MST 204.21

The Open University
Mathematics/Science/Technology
An Inter-faculty Second Level Course
MST204 Mathematical Models and Methods.

mathematical models and methods

Unit 21

Eigenvalues and eigenvectors

The Open University

Mathematics/Science/Technology
An Inter-faculty Second Level Course
MST204 Mathematical Models and Methods

Unit 21
Eigenvalues and eigenvectors

Prepared for the Course Team
by Jen Phillips

The Open University

The Open University, Walton Hall, Milton Keynes.

First published 1982. Reprinted 1983, 1985, 1989, 1991, 1992, 1994, 1996.

ISBN 0 335 14050 5

Printed and bound in the United Kingdom by Staples Printers Rochester Limited, Neptune Close, Medway City Estate, Frindsbury, Rochester, Kent ME2 4LT.

This text forms part of the correspondence element of an Open University Second Level Course.

For general availability of supporting material referred to in this text, please write to: Open University Educational Enterprises, 12 Cofferidge Close, Stony Stratford, Milton Keynes, MK11 1BY, Great Britain.

Further information on Open University courses may be obtained from The Admissions Office, The Open University, P.O. Box 48, Milton Keynes, MK7 6AB.

1.8

Contents

Introduction

In *Unit 9*, Section 5, we discussed the solution of sets of homogeneous simultaneous equations. In this unit, we examine a related problem which often arises in practice. For instance, in *Unit 24*, you will be studying systems of particles and springs like those shown in Figure 1.

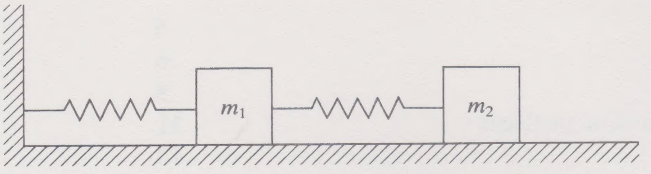

Figure 1

The equations of motion for this system are a set of linear simultaneous differential equations. In *Unit 22*, you will see that a first step in solving such a set of differential equations is to find the column vectors $\mathbf{x}$ which satisfy a set of algebraic linear equations of the form

$$\mathbf{Ax} = \lambda\mathbf{x}, \tag{1}$$

where $\mathbf{A}$ is a square matrix which is determined by the set of differential equations and λ is a number whose value is to be determined. So, before we can do the work in *Units 22* and *24*, we first have to discuss methods of solution of Equation (1). In general the only solution is $\mathbf{x} = \mathbf{0}$, but there are particular values of λ for which solutions other than $\mathbf{x} = \mathbf{0}$ are possible. These particular values of λ, and the associated non-zero column vectors $\mathbf{x}$, enable us to solve the equations of motion for the system shown in Figure 1.

Equations of the form $\mathbf{Ax} = \lambda\mathbf{x}$ also arise in a large variety of other mechanical and electrical systems in engineering. There are also statistical, numerical and other non-engineering situations in which such equations arise. In fact, they arise so frequently that it is worth discussing the problem in its own right, which is what we will do in this unit. The values of λ for which Equation (1) has non-zero solutions are called the **eigenvalues** of $\mathbf{A}$, and the non-zero solutions are called the **eigenvectors** of $\mathbf{A}$.

Study guide

You should study Section 1 and Subsections 2.1 and 2.2 before watching the television programme.

Section 4 describes a computer package which you can use to find eigenvalues and eigenvectors of a given matrix. You might like to use this package when studying *Units 22* and *24*, as well as in this unit. The exercises at the end of Section 4 are there for use at the computer terminal. You are expected to do as many as you have time for. However, you may not have time to do them all, so start with the ones that interest you the most.

The exercises at the end of Sections 1, 2 and 3 are there to give you extra practice. It is not necessarily expected that you will do them when first reading this unit. You might like to use them for revision purposes. However, you are encouraged to do the end of unit test in Section 6 immediately after studying the unit.

1 The theoretical eigenvalue problem

1.1 Introduction

In this section, we discuss a method for finding non-zero vectors **x** which satisfy the matrix equation

$$\mathbf{Ax} = \lambda\mathbf{x}. \tag{1}$$

Here, **A** is a known square matrix, but at the beginning of the problem, λ is an unknown number.

In fact, this matrix equation turns out to be a set of homogeneous linear simultaneous equations, as the following example demonstrates.

Example 1

If $\mathbf{A} = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}$, and $\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \end{bmatrix}$, show that the matrix equation $\mathbf{Ax} = \lambda\mathbf{x}$ may be written as a set of homogeneous linear simultaneous equations.

Solution

The matrix equation is

$$\begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}.$$

This gives us the equations

$$5x_1 + 2x_2 = \lambda x_1$$
$$2x_1 + 5x_2 = \lambda x_2,$$

and these can be re-written as a pair of homogeneous linear simultaneous equations

$$(5-\lambda)x_1 + \qquad 2x_2 = 0$$
$$2x_1 + (5-\lambda)x_2 = 0.$$

More generally, the equation

$$\mathbf{Ax} = \lambda\mathbf{x}$$

can be re-written as

$$\mathbf{Ax} = \lambda\mathbf{Ix},$$

I is the identity matrix defined in *Unit 20*, Subsection 4.1.

so

$$\mathbf{Ax} - \lambda\mathbf{Ix} = \mathbf{0},$$

or

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}. \tag{2}$$

The matrix equation (2) is just the matrix form of a set of homogeneous linear simultaneous equations. Such equations were discussed in Section 5 of *Unit 9*. You will remember that homogeneous sets of equations always have the 'trivial solutions' $\mathbf{x} = \mathbf{0}$. However, if the equations are linearly dependent they will have an infinite number of other solutions as well. Now, linear dependence of homogeneous sets of equations is guaranteed if the determinant of the left-hand-side coefficients is zero. So, there are non-trivial solutions to the equations

For this application of determinants see *Unit 20*, Subsection 5.3.

$$\mathbf{Ax} = \lambda\mathbf{x}$$

i.e. $\quad (\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \qquad (2)$

if λ satisfies

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0. \tag{3}$$

Example 2

The equations from Example 1:

$$\begin{aligned}(5-\lambda)x_1 + \quad\quad 2x_2 &= 0\\ 2x_1 + (5-\lambda)x_2 &= 0\end{aligned}$$

have non-trivial solutions if

$$\begin{vmatrix} 5-\lambda & 2 \\ 2 & 5-\lambda \end{vmatrix} = 0.$$

From the determinant in this example, we could have gone on to find λ.
In general, the determinant condition (3) can be used to find values of λ for which the equations have nontrivial solutions. These values of λ are called **eigenvalues** of $\mathbf{A}$.

For each eigenvalue, we get a particular set of equations, which can be solved for $\mathbf{x}$. These solutions are called **eigenvectors**, and the whole problem is often referred to loosely as the **eigenvalue problem**.

The rest of this section is made up mostly of examples and exercises. This is to help you gain some experience in solving simple eigenvalue problems.

1.2 The 2 × 2 eigenvalue problem

Example 3

Find the eigenvalues of the matrix

$$\mathbf{A} = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}.$$

Solution

The obvious way to approach this is to start with the definition of an eigenvalue, and to find the values of λ for which the equation $\mathbf{Ax} = \lambda\mathbf{x}$, that is

$$\begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \lambda\begin{bmatrix} x_1 \\ x_2 \end{bmatrix},$$

has non-trivial solutions.

We know from Example 1 that this matrix equation gives us the separate equations

$$\begin{aligned}(5-\lambda)x_1 + \quad\quad 2x_2 &= 0\\ 2x_1 + (5-\lambda)x_2 &= 0.\end{aligned}$$

For non-trivial solutions, we require the determinant of the left-hand-side coefficients to be zero. So,

$$\begin{vmatrix} 5-\lambda & 2 \\ 2 & 5-\lambda \end{vmatrix} = 0.$$

This gives us the equation

$$(5-\lambda)^2 - 2^2 = 0 \qquad (4)$$

so $\quad ((5-\lambda)-2)((5-\lambda)+2) = 0$

or $\quad (3-\lambda)(7-\lambda) = 0.$

Thus we can find two values, $\lambda = 3$ or $\lambda = 7$, for which the equations have non-trivial solutions. Hence, the eigenvalues of $\mathbf{A}$ are 3 and 7.

Each time we use the above method to find the eigenvalues of a matrix $\mathbf{A}$, we have to find the values of λ which satisfy $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. So, it is normal practice to short-cut the work shown by going directly from the matrix $\mathbf{A}$ to the determinant equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. For instance, in the example above we could easily have written down

$$\begin{vmatrix} 5-\lambda & 2 \\ 2 & 5-\lambda \end{vmatrix} = 0$$

straight away, and solved the resulting equation

$$(5-\lambda)^2 - 2^2 = 0 \tag{4}$$

as before.

The equation which we obtain by expanding the determinant (such as Equation (4) above) is often called the **characteristic equation** of the matrix.

Example 4

Find the eigenvalues of

$$\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}.$$

Solution

The condition $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ gives

$$\begin{vmatrix} 2-\lambda & 1 \\ 1 & 2-\lambda \end{vmatrix} = 0$$

and hence we obtain the characteristic equation

$$(2-\lambda)^2 - 1^2 = 0$$

i.e. $\quad (3-\lambda)(1-\lambda) = 0.$

Hence, the eigenvalues of $\mathbf{A}$ are 3 and 1.

In general, the method for finding the eigenvalues of a 2×2 matrix is as follows.

Procedure 1.2(a): To find the eigenvalues of a 2 × 2 matrix

To find the eigenvalues of

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix},$$

write down the characteristic equation $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, that is,

$$\begin{vmatrix} a_{11}-\lambda & a_{12} \\ a_{21} & a_{22}-\lambda \end{vmatrix} = 0,$$

and solve for λ.

In theory, this method can be extended to find the eigenvalues of any square matrix $\mathbf{A}$.

Exercise 1

Find the eigenvalues of the matrices

(i) $\begin{bmatrix} 7 & 3 \\ 3 & 7 \end{bmatrix}$ (ii) $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$

[*Solution on p. 45*]

Once the eigenvalues of a matrix have been found we can go on to find the corresponding eigenvectors.

Example 5

Find the eigenvectors of the matrix

$$\mathbf{A} = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}.$$

Solution

We know from Example 3 that the eigenvalues are 3 and 7. To find the corresponding eigenvectors we consider each eigenvalue in turn and solve

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0} \tag{2}$$

for $\mathbf{x}$.

Case $\lambda = 3$: in this case Equation (2) becomes

$$\begin{aligned} (5-3)x_1 + \quad\quad 2x_2 &= 0 \\ 2x_1 + (5-3)x_2 &= 0. \end{aligned}$$

The fact that both these equations give us the same information

$$x_1 + x_2 = 0$$

is not surprising, as the determinant condition forces the equations to be linearly dependent.

To solve these equations we put $x_2 = k$, where k is an arbitrary number. Substituting this into the equation above, we get $x_1 = -k$. In this way we obtain an infinite number of solutions of the form

This approach is first discussed in Section 2 of *Unit 9*.

$$\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \begin{bmatrix} -k \\ k \end{bmatrix}.$$

So, with $\lambda = 3$, any vector of the form

$$\mathbf{x} = k\begin{bmatrix} -1 \\ 1 \end{bmatrix}$$

will satisfy Equation (2).

When we give an eigenvector corresponding to a given eigenvalue, it is common practice to omit the k, and just to say that an eigenvector corresponding to the eigenvalue 3 is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$. In doing this, it is understood that $\begin{bmatrix} 1 \\ -1 \end{bmatrix}$, $\begin{bmatrix} -2 \\ 2 \end{bmatrix}$, etc. would be equally correct, and that an eigenvector is not unique.

Case $\lambda = 7$: in this case Equation (2) becomes

$$\begin{aligned} (5-7)x_1 + \quad\quad 2x_2 &= 0 \\ 2x_1 + (5-7)x_2 &= 0. \end{aligned}$$

These give us the single equation

$$-x_1 + x_2 = 0.$$

Putting $x_2 = k$ into this equation, we get $x_1 = k$. Thus, any vector of the form $\begin{bmatrix} k \\ k \end{bmatrix}$ is a solution of $\mathbf{A}\mathbf{x} = 7\mathbf{x}$ and so an eigenvector corresponding to the eigenvalue 7 is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

Procedure 1.2(b): To find the eigenvectors of a 2 × 2 matrix

To find the eigenvectors of

$$\mathbf{A} = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}:$$

1. Find the eigenvalues λ_1 and λ_2 of $\mathbf{A}$.

2. For each of these values of λ solve $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$ for $\mathbf{x}$. That is, solve

$$\left.\begin{aligned} (a_{11} - \lambda_k)x_1 + \quad\quad a_{12}x_2 &= 0 \\ a_{21}x_1 + (a_{22} - \lambda_k)x_2 &= 0 \end{aligned}\right\} \text{for } k = 1, 2.$$

Then any non-trivial solution to these equations will give an eigenvector corresponding to the eigenvalue λ_k.

Again, the method can *in theory* be extended to find the eigenvectors of any square matrix $\mathbf{A}$.

Exercise 2

Use the eigenvalues you found in Exercise 1 to find eigenvectors of

(i) $\begin{bmatrix} 7 & 3 \\ 3 & 7 \end{bmatrix}$, (ii) $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$.

[*Solution on p. 45*]

1.3 The 3 × 3 eigenvalue problem

The problem of finding eigenvalues and eigenvectors of a 3 × 3 matrix can be tackled in the same way as for a 2 × 2 matrix, except that the algebra on the way is likely to be more complicated.

Example 6

Find the eigenvalues and corresponding eigenvectors of the matrix

$$\mathbf{A} = \begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 5 \end{bmatrix}$$

Solution

First, we find the eigenvalues, by finding the values of λ which satisfy $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$, that is

$$\begin{vmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 0 & 0 & 5-\lambda \end{vmatrix} = 0.$$

This determinant can be evaluated in various ways. One way is to interchange rows 1 and 3, to obtain a form which is easier to expand:

$$-\begin{vmatrix} 0 & 0 & 5-\lambda \\ 1 & 2-\lambda & 1 \\ 2-\lambda & 1 & 1 \end{vmatrix} = 0.$$

Here we have used the property that interchanging two rows of a determinant changes its sign. (*Unit 20*, Section 5, Property 1.)

Now, expanding by the top row, we obtain the characteristic equation

$$-(5-\lambda)(1-(2-\lambda)^2) = 0$$

or $\quad (5-\lambda)(3-\lambda)(1-\lambda) = 0.$

So the eigenvalues are 5, 3 and 1.

To find the eigenvectors, we solve $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, that is

$$\begin{bmatrix} 2-\lambda & 1 & 1 \\ 1 & 2-\lambda & 1 \\ 0 & 0 & 5-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}, \tag{5}$$

putting λ equal to each eigenvalue in turn.

Case $\lambda = 5$: in this case Equation (5) becomes

$$\begin{aligned} -3x_1 + x_2 + x_3 &= 0 \qquad E_1 \\ x_1 - 3x_2 + x_3 &= 0. \qquad E_2 \end{aligned}$$

Note in this case that the third equation is $0 = 0$. Since there are now effectively 2 equations in 3 unknowns there is an infinite number of solutions.

We can find the solution to these equations, using Gaussian elimination. $E_2 + \frac{1}{3}E_1$ gives

Gaussian elimination was discussed in *Unit 9*.

$$-\tfrac{8}{3}x_2 + \tfrac{4}{3}x_3 = 0. \qquad E_{2a}$$

Now putting $x_3 = k$, from E_{2a} we get $x_2 = \dfrac{k}{2}$. Substituting these values into E_1, we get

$$-3x_1 + \frac{k}{2} + k = 0$$

which gives $x_1 = \frac{k}{2}$. Thus we get the solution to the equations:

$$x_1 = \frac{k}{2}, \qquad x_2 = \frac{k}{2}, \qquad x_3 = k.$$

Hence, an eigenvector corresponding to the eigenvalue 5 is

$$\begin{bmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 1 \end{bmatrix}.$$

To economize on space we shall often print column vectors such as this in the form $[\frac{1}{2} \quad \frac{1}{2} \quad 1]^T$, using the transpose notation introduced in *Unit 20.*

An equally good eigenvector would be $[1 \quad 1 \quad 2]^T$.

Case $\lambda = 3$: putting this value of λ into Equation (5), we obtain

$$\begin{aligned} -x_1 + x_2 + x_3 &= 0 \qquad E_1 \\ x_1 - x_2 + x_3 &= 0 \qquad E_2 \\ 2x_3 &= 0 \qquad E_3. \end{aligned}$$

From E_3, we see that $x_3 = 0$. Substituting $x_3 = 0$ into E_1 and E_2 we see that they both provide us with the same information,

$$x_1 - x_2 = 0,$$

from which we obtain a solution

$$x_1 = k, \qquad x_2 = k, \qquad x_3 = 0$$

and hence $[1 \quad 1 \quad 0]^T$ is an eigenvector corresponding to the eigenvalue 3.

Case $\lambda = 1$: finally, putting $\lambda = 1$ into Equation (5), we obtain the equations

$$\begin{aligned} x_1 + x_2 + x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \\ 4x_3 &= 0. \end{aligned}$$

Again, from the last equation, $x_3 = 0$. Substituting this into the first two equations, we obtain the single equation

$$x_1 + x_2 = 0.$$

Hence

$$x_1 = k, \qquad x_2 = -k, \qquad x_3 = 0$$

is a solution to the equations, and $[1 \quad -1 \quad 0]^T$ is an eigenvector corresponding to the eigenvalue 1.

Exercise 3

Find the eigenvalues and corresponding eigenvectors of the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 6 \\ 0 & 4 & 4 \\ 0 & 0 & 2 \end{bmatrix}.$$

[*Solution on p. 45*]

For a 3×3 matrix, the characteristic equation which we get from $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ will be a cubic. In all the examples and exercises which we have used in this section, the characterstic equation factorizes easily, and has integer roots, but this is hardly likely to be the case in practical situations.

In general, you will have to use some numerical method to find the eigenvalues. If you have access to a computer (or a programmable calculator), then the methods described in the rest of this unit should be used. If not, then you could always find the roots of the characteristic equation by using the Newton–Raphson method

described in *Unit 18*. Sketch the graph of the characteristic polynomial, and use the approximations you find from the graph as the three starting values for Newton–Raphson. In this way, find each of the two roots to the accuracy of your choice. After this, the eigenvectors can be found in much the same way as before.

You could also use the methods described in the next section to find these eigenvalues—there is not a great deal of difference in the amount of work involved.

In either case, this approach is only suitable if the roots of the characteristic equation are real. Fortunately, in practical applications, like the ones you will meet in *Unit 24*, they usually are.

For anything larger than a 3×3 matrix, the whole process becomes extremely laborious, and you would be advised to get a computer package to help you.

1.4 Nature of solutions to the eigenvalue problem

So far in this section, all the matrices have had *real* eigenvalues. For instance, the 3×3 matrix in Example 6

$$\begin{bmatrix} 2 & 1 & 1 \\ 1 & 2 & 1 \\ 0 & 0 & 5 \end{bmatrix} \tag{6}$$

has three real eigenvalues 5, 3 and 1.

However, there are many matrices which do not have this property. For instance, the matrix in the example below has two complex eigenvalues.

Example 7

The matrix $\mathbf{B} = \begin{bmatrix} 1 & -1 \\ 4 & 1 \end{bmatrix}$ has eigenvalues given by

$$\begin{vmatrix} 1-\lambda & -1 \\ 4 & 1-\lambda \end{vmatrix} = 0.$$

The characteristic equation

$$(1-\lambda)^2 + 4 = 0$$

can be simplified (using $(1-\lambda)^2 + 4 = (1-\lambda)^2 - (2i)^2$) to give

$$(1-\lambda+2i)(1-\lambda-2i) = 0.$$

Thus, $\mathbf{B}$ has complex eigenvalues $1 + 2i$ and $1 - 2i$.

Exercise 4

Find the eigenvectors of the matrix in Example 7.

[*Solution on p. 45*]

Another important property which the matrices we have looked at so far have had, is that all the 2×2 matrices have had *two* eigenvalues, and all the 3×3 matrices have had *three* eigenvalues. For instance, the matrix (6) above has three distinct eigenvalues 5, 3 and 1. More generally, it is usually the case that an $n \times n$ matrix has n distinct eigenvalues. However, there are many matrices which do not have this property, as demonstrated in the following example.

Example 8

The matrix $\mathbf{A} = \begin{bmatrix} 5 & 3 \\ -3 & -1 \end{bmatrix}$ has eigenvalues given by

$$\begin{vmatrix} 5-\lambda & 3 \\ -3 & -1-\lambda \end{vmatrix} = 0.$$

The characteristic equation

$$(5-\lambda)(-1-\lambda) + 9 = 0$$

can be simplified to give

$$\lambda^2 - 4\lambda + 4 = 0$$

or $$(\lambda - 2)^2 = 0.$$

In this case, we say that $\mathbf{A}$ has a **repeated** eigenvalue of 2, and that the eigenvalues of $\mathbf{A}$ are not **distinct**.

When we try to calculate the eigenvectors we obtain the equation

$$x_1 + x_2 = 0.$$

This equation has a solution of the form $x_1 = k$, $x_2 = -k$, and hence there is an eigenvector $[1 \quad 1]^T$ corresponding to the eigenvalue 2.

Note that the matrix in Example 8 effectively has only one eigenvector corresponding to the repeated eigenvalue 2, whereas all the other 2×2 matrices we have met have had two distinct eigenvalues and eigenvectors. However, there are cases in which a repeated eigenvalue gives rise to more than one eigenvector. This is illustrated by the following example.

Example 9

By inspection, it can be seen that the eigenvalues of

$$\mathbf{A} = \begin{bmatrix} 2 & 0 & 1 \\ 0 & 2 & 1 \\ 0 & 0 & 3 \end{bmatrix}$$

are given by

$$(2 - \lambda)^2(3 - \lambda) = 0.$$

Thus, $\mathbf{A}$ has two eigenvalues 2 and 3. The eigenvalue 2 is a repeated eigenvalue.

Putting $\lambda = 3$ into $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, we obtain the equations

$$-x_1 + x_3 = 0$$
$$-x_2 + x_3 = 0.$$

Hence we get an eigenvector $[1 \quad 1 \quad 1]^T$ corresponding to the eigenvalue 3.

Putting $\lambda = 2$ into $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, we obtain only the information

$$x_3 = 0.$$

So x_1 and x_2 can assume any values we choose, and these values are not related to each other. Thus, any vector of the form $[k \quad l \quad 0]^T$ will be an eigenvector. In particular $[1 \quad 0 \quad 0]^T$ and $[0 \quad 1 \quad 0]^T$ are both eigenvectors, and these are essentially distinct. More precisely we mean that the vectors are linearly independent. Further, the third eigenvector $[1 \quad 1 \quad 1]^T$ is essentially distinct from both these in the sense that all three vectors are linearly independent. This can be shown by forming the vectors into the columns of a matrix $\mathbf{P}$, say

$$\mathbf{P} = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 1 \\ 0 & 0 & 1 \end{bmatrix}.$$

Then $\det \mathbf{P} = 1$, which is non-zero, and this is sufficient to show that the rows and columns of $\mathbf{P}$ are linearly independent.

See *Unit 20* Subsection 4.4.

In general, an $n \times n$ matrix has *at most* n distinct eigenvalues. If there are exactly n there will always be n linearly independent eigenvectors, but if any of the eigenvalues are repeated there may or may not be n linearly independent eigenvectors, as illustrated by Examples 8 and 9.

In the numerical methods we discuss later in this unit, it is assumed that the matrix in question has n real and distinct eigenvalues. In fact, a lot of matrices which occur in practice do have this property. In particular, symmetric matrices always have real eigenvalues and a full set of linearly independent eigenvectors.

A symmetric matrix $\mathbf{A}$ is one for which $\mathbf{A}^T = \mathbf{A}$ (see *Unit 20*, Subsection 4.4).

Summary of Section 1

The set of homogeneous linear simultaneous equations given by

$$\mathbf{Ax} = \lambda\mathbf{x}, \qquad (1)$$

where $\mathbf{A}$ is a known square matrix, has non-trivial solutions for various values of λ. These values of λ are called the **eigenvalues** of $\mathbf{A}$. For each eigenvalue λ there corresponds an infinite set of solutions to Equation (1). Any non-zero solution $\mathbf{x}$ in this set is called an **eigenvector**.

One way of finding the eigenvalues of $\mathbf{A}$ is to find the values of λ which satisfy the characteristic equation

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

To find an eigenvector corresponding to an eigenvalue λ, find *any* non-zero solution to the equation

$$(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}.$$

The resulting eigenvector $\mathbf{x}$ is not unique.

End of section exercise

Exercise 5

Find the eigenvalues and the corresponding eigenvectors for the following matrices.

(i) $\begin{bmatrix} 1 & 2 \\ -1 & 4 \end{bmatrix}$ (ii) $\begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix}$ (iii) $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$

(iv) $\begin{bmatrix} 1 & 0 & -1 \\ 1 & 2 & 1 \\ 2 & 2 & 3 \end{bmatrix}$ (v) $\begin{bmatrix} 3 & 2 & 2 \\ 2 & 2 & 0 \\ 2 & 0 & 4 \end{bmatrix}$

Hint: if you get stuck, (iv) has an eigenvalue 1.

[*Solution on p. 46*]

2 Iterative methods for finding selected eigenvalues (Television Section)

2.1 Eigenvalues of matrices related to a given matrix A

The television programme can roughly be divided into two parts. The first part of the programme explains the geometric significance of the eigenvalues and eigenvectors of a 2×2 matrix. The second part of the programme makes use of an observed geometric property of eigenvectors to construct some numerical methods for finding particular eigenvalues and eigenvectors.

To understand what is going on in the second part of the programme, you will need to know how the eigenvalues (and eigenvectors) of $\mathbf{A}^{-1}$, and $(\mathbf{A} - p\mathbf{I})^{-1}$ are related to those of $\mathbf{A}$. For the purposes of this subsection, $\mathbf{A}$ is an $n \times n$ matrix with n real distinct eigenvectors, We shall call these

$$\lambda_1, \lambda_2, \ldots, \lambda_n.$$

Since the eigenvalues are distinct, we know from Subsection 1.4 that there must be a corresponding set of linearly independent eigenvectors

$$\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n,$$

where the $\mathbf{x}$s are non-zero.

We know that the eigenvalues and eigenvectors of $\mathbf{A}$ are related by the equations

$$\mathbf{Ax}_1 = \lambda_1\mathbf{x}_1, \qquad \mathbf{Ax}_2 = \lambda_2\mathbf{x}_2, \qquad \ldots, \qquad \mathbf{Ax}_n = \lambda_n\mathbf{x}_n.$$

More briefly we can express all these equations as

$$\mathbf{Ax}_i = \lambda_i\mathbf{x}_i \qquad \text{for } i = 1, 2, .., n. \qquad (1)$$

We can use these equations to find the relationship between the eigenvalues of $\mathbf{A}$ and $\mathbf{A}^2$, as the following example demonstrates.

Example 1

Show that the eigenvalues of $\mathbf{A}^2$ are $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$ and that the eigenvectors are the same as those of $\mathbf{A}$.

Solution

We know from (1) that if $\mathbf{A}$ has a set of eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, then

$$\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n \tag{1}$$

where $\mathbf{x}_i$ is an eigenvector corresponding to λ_i.

Suppose, for each i, we left multiply both sides of (1) by $\mathbf{A}$. Then we obtain

$$\begin{aligned} \mathbf{A}^2\mathbf{x}_i &= \mathbf{A}(\lambda_i \mathbf{x}_i) \\ &= \lambda_i(\mathbf{A}\mathbf{x}_i) \qquad &&(\text{as } \lambda_i \text{ is just a number}) \\ &= \lambda_i(\lambda_i \mathbf{x}_i) \qquad &&(\text{from (1)}) \end{aligned}$$

so

$$\mathbf{A}^2\mathbf{x}_i = \lambda_i^2 \mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n. \tag{2}$$

This last equation tells us that $\mathbf{A}^2$ has eigenvalues $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$, with corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$—which are the same as those of $\mathbf{A}$.

So, we conclude that

(i) $\mathbf{A}^2$ has eigenvalues $\lambda_1^2, \lambda_2^2, \ldots, \lambda_n^2$;

(ii) $\mathbf{A}$ and $\mathbf{A}^2$ have the same eigenvectors.

The eigenvalues of $\mathbf{A} + q\mathbf{I}$

Using an approach similar to that used in Example 1 we can find the eigenvalues of $\mathbf{A} + q\mathbf{I}$, for any number q.

We know that

$$\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n. \tag{1}$$

To obtain the eigenvalues of $\mathbf{A} + q\mathbf{I}$, we add $q\mathbf{x}_i$ (or equivalently $q\mathbf{I}\mathbf{x}_i$) to both sides of (1). We get

$$\mathbf{A}\mathbf{x}_i + q\mathbf{I}\mathbf{x}_i = \lambda_i \mathbf{x}_i + q\mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n$$

which can be rewritten as

$$(\mathbf{A} + q\mathbf{I})\mathbf{x}_i = (\lambda_i + q)\mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n. \tag{3}$$

You will need this result in Subsection 3.3.

So, we conclude that $\mathbf{A} + q\mathbf{I}$ has n distinct eigenvalues $\lambda_1 + q, \quad \lambda_2 + q, \quad \ldots, \quad \lambda_n + q$, with corresponding eigenvectors $x_1, x_2, \ldots, x_n$—which are the same as those of $\mathbf{A}$.

Exercise 1

Given that $\mathbf{A}$ has a set of eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$, state the eigenvalues of (i) $\mathbf{A} + 5\mathbf{I}$, (ii) $\mathbf{A} - 2\mathbf{I}$, (iii) $\mathbf{A} - q\mathbf{I}$.

[*Solution on p. 47*]

The eigenvalues of $\mathbf{A}^{-1}$

Assuming that $\mathbf{A}$ is non-singular we can use the approach of Example 1 to find the eigenvalues of $\mathbf{A}^{-1}$. Again we start with Equation (1):

$$\mathbf{A}\mathbf{x}_i = \lambda_i \mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots n.$$

Multiplying by $\mathbf{A}^{-1}$, we get

$$\mathbf{x}_i = \lambda_i \mathbf{A}^{-1}\mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n$$

and dividing by λ_i, we obtain

Note: $\lambda_i \neq 0$, for otherwise $\mathbf{x}_i$ $(= \lambda_i \mathbf{A}^{-1}\mathbf{x}_i)$ would be zero, contrary to the definition of 'eigenvector'.

$$\mathbf{A}^{-1}\mathbf{x}_i = \frac{1}{\lambda_i}\mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n. \tag{4}$$

From Equation (4), we conclude that $\mathbf{A}^{-1}$ has n distinct eigenvalues $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_n}$, with corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$—which are the same as those of $\mathbf{A}$.

This result is needed for the television programme.

The eigenvalues of $(\mathbf{A} - p\mathbf{I})^{-1}$

In the television programme we will need to know the eigenvalues and eigenvectors of matrices of the form $(\mathbf{A} - p\mathbf{I})^{-1}$ for various values of p. The following exercise asks you to express these eigenvalues and eigenvectors in terms of those of $\mathbf{A}$.

Exercise 2

Given that $\mathbf{A}$ has n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ with corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, and that p is a number such that $\mathbf{A} - p\mathbf{I}$ is a non-singular matrix, show that:

(i) $(\mathbf{A} - p\mathbf{I})^{-1}$ has n distinct eigenvalues

$$\frac{1}{\lambda_1 - p}, \quad \frac{1}{\lambda_2 - p}, \quad \ldots, \quad \frac{1}{\lambda_n - p};$$

(ii) the corresponding eigenvectors are the same as those of $\mathbf{A}$.

[*Solution on p. 47*]

2.2 Pre-television notes

The following facts are assumed in the television programme.

1. The product of a matrix and a column vector gives another column vector. For instance,

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} = \begin{bmatrix} 11 \\ 7 \end{bmatrix}.$$

Unit 20, Section 2.4.

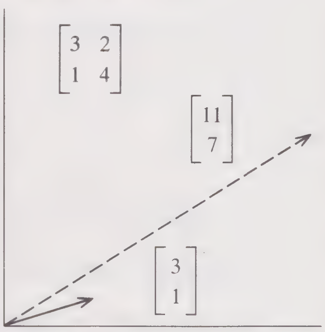

Figure 1

In the last unit we used matrix multiplication to describe a change of co-ordinate axes. The geometrical interpretation used in this unit is slightly different, in that we *fix* the co-ordinate axes and say that a matrix acting on a vector produces a new vector. The process of producing a new vector in this way is often referred to as a **transformation**. For example $[3 \quad 1]^T$ is moved to $[11 \quad 7]^T$ by the transformation described by the matrix $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ as in Figure 1. In particular, the television programme refers to transformations.

2. You will remember, from Section 1, that the eigenvalues of a 2×2 matrix $\mathbf{A}$ can be found by solving

$$\det(\mathbf{A} - \lambda\mathbf{I}) = 0.$$

As this invariably produces a quadratic equation, there are three possibilities:

(i) $\mathbf{A}$ has two distinct real eigenvalues;

(ii) $\mathbf{A}$ has two equal real eigenvalues;

(iii) $\mathbf{A}$ has two complex eigenvalues.

The matrices referred to in the television programme are $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ and $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$.

In Section 1 you found that $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ has 2 distinct real eigenvalues with eigenvectors $[1 \quad 1]^T$ and $[1 \quad -\frac{1}{2}]^T$. If you have done Exercise 4 at the end of Section 1 you will have found that the second matrix $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ has complex eigenvalues.

3. From the work in Subsection 2.1, you will need to know that if $\mathbf{A}$ has the set of eigenvalues $\{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ with corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, then

(i) $\mathbf{A}^{-1}$ has the set of eigenvalues

$$\left\{\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_n}\right\}.$$

so long as $\mathbf{A}$ is non-singular.

(ii) $(\mathbf{A} - p\mathbf{I})^{-1}$ has the set of eigenvalues

$$\left\{\frac{1}{\lambda_1 - p}, \frac{1}{\lambda_2 - p}, \ldots, \frac{1}{\lambda_n - p}\right\},$$

so long as $\mathbf{A} - p\mathbf{I}$ is non-singular.

(iii) $\mathbf{A}$, $\mathbf{A}^{-1}$ and $(\mathbf{A} - p\mathbf{I})^{-1}$ all have the same eigenvectors.

Now watch the television programme. In the second part of the programme, concentrate on the methods, rather than the arithmetical details. **TV 21**

The following subsections all contain references to the television programme. Specific references will be marked with a [TV] *symbol.*

2.3 Geometric interpretation of eigenvalues and eigenvectors

(*Summary of first part of the television programme*)

TV

1. The programme examines the effect that the matrix $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ has on different vectors. In this way a vector $\mathbf{x}$ is found for which the vector formed by taking

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}\mathbf{x}$$

is in the *same* direction as $\mathbf{x}$.

In fact, it is shown that a transformation on *any* vector in this 'special direction':

(i) leaves the direction unchanged

(ii) stretches the vector to five times its original length, that is

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}\mathbf{x} = 5\mathbf{x}. \qquad \text{(See Figure 2.)}$$

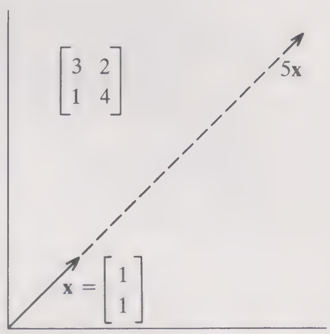

Figure 2

2. Other matrices are examined to see if they have associated 'special directions'. It is found that some do, and some do not. For intance, the matrix $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ has no associated 'special direction'. It is, in fact, a multiple of the rotation matrix

$$\begin{bmatrix} \frac{1}{\sqrt{2}} & -\frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} & \frac{1}{\sqrt{2}} \end{bmatrix} = \begin{bmatrix} \cos\frac{\pi}{4} & -\sin\frac{\pi}{4} \\ \sin\frac{\pi}{4} & \cos\frac{\pi}{4} \end{bmatrix},$$

and has the effect of rotating every vector it operates on through an angle $\frac{\pi}{4}$ anti-clockwise, and enlarging it.

3. To discover whether it is possible to predict which matrices have 'special directions' and, if they do, by how much the vectors in this direction are stretched, the programme looks at the algebra of what has been happening. We have been looking for 'special vectors' $\mathbf{x}$ such that

$$\mathbf{A}\mathbf{x} = \lambda\mathbf{x},$$

where λ is the stretch factor. This equation is just the algebraic eigenvalue problem discussed in the first section.

Exercise 5 of Section 1 shows that the eigenvalues and eigenvectors of $\begin{bmatrix} 1 & -1 \\ 1 & 1 \end{bmatrix}$ are complex. For a geometric interpretation, however, we need a real eigenvector. As this matrix has no real eigenvectors, it has no special direction either.

In the pre-television notes we recalled that $\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$ has two eigenvectors: $[1 \quad 1]^T$ and $[1 \quad -\frac{1}{2}]^T$. The television programme checks geometrically that $[1 \quad -\frac{1}{2}]^T$ is a special direction, and that the enlargement in this direction is indeed the corresponding eigenvalue.

The conclusion we come to is that for some 2×2 matrices $\mathbf{A}$, there are 'special vectors' $\mathbf{x}$, which when transformed do not change their direction. These special vectors, when they exist, are eigenvectors of $\mathbf{A}$. The factor by which each of these is stretched can be found by calculating the equivalent eigenvalue.

So, if $\mathbf{A}$ has real eigenvalues λ with corresponding eigenvectors $\mathbf{x}$, then

(i) $\mathbf{x}$ will be in a 'special direction'

(ii) λ will be the amount by which $\mathbf{x}$ is stretched.

2.4 Direct iteration

The television programme takes a matrix $\mathbf{A}$ which is known to have two real eigenvectors. Arbitrary vectors $\mathbf{y}$ are taken, and transformed by $\mathbf{A}$ to give new vectors $\mathbf{Ay}$. All appear to be closer than the corresponding vector $\mathbf{y}$ to one of the 'special directions' (see Figure 3). This holds provided that $\mathbf{y}$ was not in a special direction to start with, for then $\mathbf{Ay}$ is known to be in the same direction as $\mathbf{y}$. To see why the new vector is closer to the special direction than the old one let us look at the matrix $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$, which we know has eigenvalues 5 and 2 with corresponding eigenvectors $[1 \quad 1]^T$ and $[1 \quad -\frac{1}{2}]^T$.

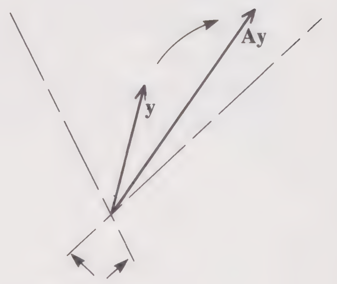

Figure 3

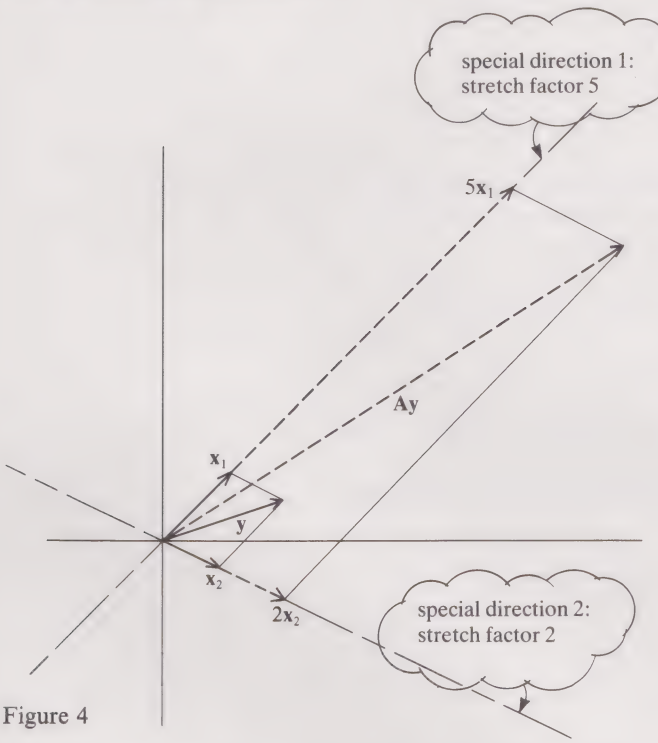

Figure 4

Any two-dimensional vector $\mathbf{y}$ not in one of the special directions can be expressed as the sum of vectors in the two special directions. As vectors in these directions are eigenvectors, $\mathbf{y}$ can simply be expressed as the sum of two eigenvectors of suitable lengths:

$$\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2. \tag{5}$$

Remember, eigenvectors are not unique.

Multiplying (5) by $\mathbf{A}$, we find that

$$\begin{aligned} \mathbf{Ay} &= \mathbf{Ax}_1 + \mathbf{Ax}_2 \\ &= 5\mathbf{x}_1 + 2\mathbf{x}_2. \end{aligned}$$

$\mathbf{Ax}_1 = 5\mathbf{x}_1$ and $\mathbf{Ax}_2 = 2\mathbf{x}_2$ as $\mathbf{x}_1$ and $\mathbf{x}_2$ are eigenvectors.

Now the stretching in the $\mathbf{x}_1$ direction is greater than that in the $\mathbf{x}_2$ direction. So $\mathbf{Ay}$ is closer than the original vector $\mathbf{y}$ to the $\mathbf{x}_1$ direction.

More generally we can consider any 2×2 matrix with distinct real eigenvalues λ_1 and λ_2. For then an arbitrary vector $\mathbf{y}$ (not itself an eigenvector) can be expressed as the sum of two eigenvectors:

$$\mathbf{y} = \mathbf{x}_1 + \mathbf{x}_2. \tag{5}$$

Hence

$$\mathbf{Ay} = \mathbf{Ax}_1 + \mathbf{Ax}_2$$
$$= \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2 \qquad \text{(since } \mathbf{x}_1 \text{ and } \mathbf{x}_2 \text{ are eigenvectors).}$$

Now the vector $\mathbf{Ay}$ is closer than $\mathbf{y}$ to one of the special directions specified by $\mathbf{x}_1$ and $\mathbf{x}_2$. Which one it is closer to depends on which of λ_1 and λ_2 has the larger magnitude. In general, $\mathbf{Ay}$ will be closer than $\mathbf{y}$ to the eigenvector which corresponds to the eigenvalue of *largest* modulus.

This vector is referred to in the television programme as the **dominant eigenvector**. The corresponding eigenvalue is called the **dominant eigenvalue**.

Provided that the eigenvalues are real, there are only two exceptions to this statement:

(i) if $\mathbf{y}$ happens to be an eigenvector, then $\mathbf{Ay}$ must be in the same direction as $\mathbf{y}$

(ii) if $\mathbf{A}$ has two eigenvalues of equal modulus, then the statement fails because there is no dominant eigenvector.

However, the fact that, in general, $\mathbf{Ay}$ is closer than $\mathbf{y}$ to the dominant eigenvector means that if we multiply the vector $\mathbf{Ay}$ by $\mathbf{A}$ again, the new vector formed will be even nearer to the dominant eigenvector. This observation gives rise to an iterative method for finding the dominant eigenvector. If we start with a vector $\mathbf{y}_0$ (not an eigenvector), and form the sequence of vectors

$$\begin{aligned} \mathbf{y}_1 &= \mathbf{Ay}_0 \\ \mathbf{y}_2 &= \mathbf{Ay}_1 \\ &\vdots \\ \mathbf{y}_{r+1} &= \mathbf{Ay}_r \end{aligned}$$

then each successive vector will be nearer than the previous one to an eigenvector corresponding to the eigenvalue of largest modulus. If we continue in this way then, for large enough r, $\mathbf{y}_r$ will effectively be an eigenvector of $\mathbf{A}$.

This process is called **direct iteration**, and is demonstrated in the following example.

Example 2

Given a matrix $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}$, and an arbitrary non-zero vector $\mathbf{y}_0 = [1 \quad 1]^T$ (not an eigenvector of the matrix), then the iterative scheme above can be used to find an estimate for the dominant eigenvector.

$$\mathbf{y}_1 = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}$$

$$\mathbf{y}_2 = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 7 \end{bmatrix} = \begin{bmatrix} 29 \\ 43 \end{bmatrix}$$

$$\mathbf{y}_3 = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 29 \\ 43 \end{bmatrix} = \begin{bmatrix} 173 \\ 259 \end{bmatrix}$$

and so on.

To see how quickly the process is converging, we can check the ratios of the elements. We can do this by dividing by the element of largest modulus. Thus dividing $\mathbf{y}_2$ by 43 gives $\mathbf{y}_2 = 43\,[0.67 \quad 1]^T$. Similarly, dividing $\mathbf{y}_3$ by 259 gives $\mathbf{y}_3 = 259\,[0.67 \quad 1]^T$, so, it seems likely that a required eigenvector of $\mathbf{A}$ is approximately $[0.67 \quad 1]^T$.

It is also possible to find the dominant eigenvalue using this process—but first we shall amend the process into a more usable form. You may have noticed that the numbers in the example above were getting larger and larger. This is more than a mere inconvenience. If the numbers were allowed to go on growing in this way, it is possible for the numbers in the successive column vectors $\mathbf{y}_r$ to get too large even for a computer. However, as it is only the ratio of the numbers in the vector that matter, we can scale each vector at each step so that the elements in the vectors stay a reasonable size. There are many ways of scaling, but the one usually used is to divide by the element of largest modulus at each stage of the iteration.

For instance, if at some stage the iteration gave the vector $\mathbf{y}' = [2 \quad -1]^T$ we would divide by 2 to obtain $\mathbf{y} = [1 \quad -\frac{1}{2}]^T$. Similarly, if the iteration gave the vector $\mathbf{z}' = [1 \quad -3]^T$ we would divide by -3 to obtain $\mathbf{z} = [-\frac{1}{3} \quad 1]^T$. In this way one of the elements in the scaled vector will be 1 and none of the other elements will be greater than 1. This is demonstrated geometrically in the television programme, by showing that each 2-dimensional vector formed is in turn scaled onto a point on the square (see Figure 5).

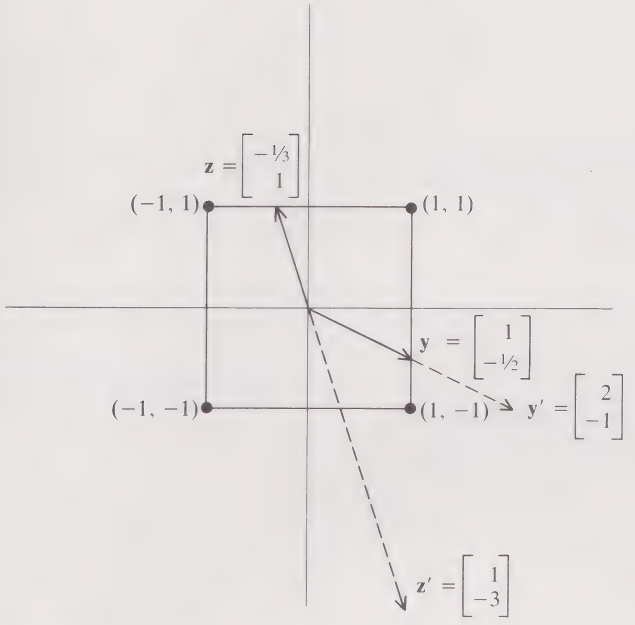

Figure 5

The following example demonstrates this scaling process. The notation $\mathbf{y}_r'$ is adopted for the unscaled vector, and $\mathbf{y}_r$ for the vector after it has been scaled.

Example 3

Re-working the last example, but this time dividing through by the element of largest modulus at each stage, we get a new sequence of vectors:

All results were worked to full accuracy, but are recorded here to three decimal places for convenience.

(i) $\mathbf{y}_1' = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ 7 \end{bmatrix}.$

Dividing by 7 gives $\mathbf{y}_1 = [0.714 \quad 1]^T$.

(ii) $\mathbf{y}_2' = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 0.714 \\ 1 \end{bmatrix} = \begin{bmatrix} 4.143 \\ 6.143 \end{bmatrix}.$

Dividing by 6.143 gives $\mathbf{y}_2 = [0.674 \quad 1]^T$.

(iii) $\mathbf{y}_3' = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 0.674 \\ 1 \end{bmatrix} = \begin{bmatrix} 4.023 \\ 6.023 \end{bmatrix}.$

Dividing by 6.023 gives $\mathbf{y}_3 = [0.668 \quad 1]^T$.

This time, the numbers have stayed under control. To see that $[0.67 \quad 1]^T$ is a reasonable approximation for the dominant eigenvector, look at $\mathbf{y}_2$ and $\mathbf{y}_3$; these are both $[0.67 \quad 1]^T$ to two significant figures.

In general, to see if a required degree of convergence has been reached, we check that each element of $\mathbf{y}_{r+1} - \mathbf{y}_r$ is small enough in magnitude to satisfy the accuracy we require. For instance, in the example above $\mathbf{y}_3 - \mathbf{y}_2 = [-0.006 \quad 0]^T$, so the largest magnitude is 0.006 which is slightly larger than the value 0.005 required for two significant figure accuracy. In practice therefore, we really should do another iteration—and check that the largest magnitude in $\mathbf{y}_4 - \mathbf{y}_3$ is less than 0.005 before we can be satisfied that the given answer is indeed correct to two significant figures. The dominant eigenvector is actually $[\frac{2}{3} \quad 1]^T$.

This process also gives us an approximation for the dominant eigenvalue. From Part (iii) of the iteration, we know that

$$\mathbf{y}_3 = \frac{\mathbf{y}'_3}{6.023} = \frac{\mathbf{A}\mathbf{y}_2}{6.023},$$

and, since $\mathbf{y}_2 \simeq \mathbf{y}_3$, we have

$$\mathbf{y}_3 \simeq \frac{\mathbf{A}\mathbf{y}_3}{6.023}.$$

This can be re-written as

$$\mathbf{A}\mathbf{y}_3 \simeq 6.023\mathbf{y}_3,$$

which tells us that as well as $\mathbf{y}_3$ being an approximation for the dominant eigenvector, 6.023 must be an approximation for the dominant eigenvalue.

In this case, three iterations have produced a good approximation.

Thus we obtain the following iterative scheme.

Procedure 2.4: Direct iteration

To find the eigenvalue of largest modulus and corresponding eigenvector of a given matrix **A**, start with an arbitrary non-zero vector $\mathbf{y}_0$ (not an eigenvector) and then

1. form $\mathbf{y}'_{r+1} = \mathbf{A}\mathbf{y}_r$,
2. form $\mathbf{y}_{r+1} = \dfrac{\mathbf{y}'_{r+1}}{\alpha_{r+1}}$,

where α_{r+1} is the element of largest modulus in $\mathbf{y}'_{r+1}$.

Do 1 and 2 for $r = 0, 1, 2, \ldots$.

Then, for sufficiently large r, α_{r+1} will be a good approximation to the eigenvalue of largest modulus of **A**, and a corresponding eigenvector will be approximately equal to $\mathbf{y}_{r+1}$.

This is often referred to as the **power method**.

This scheme will work, so long as

(i) the eigenvalues are real,

(ii) the moduli of the eigenvalues are distinct, and

(iii) the initial vector $\mathbf{y}_0$ is not an eigenvector.

Exercise 3

Use the procedure above to find $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_3$ for the matrix

$$\mathbf{A} = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix}$$

starting with $\mathbf{y}_0 = [1 \quad 1]^T$.

Hence, estimate the eigenvalue of largest modulus, and a corresponding eigenvector.

[*Solution on p. 47*]

This procedure for finding the eigenvalue of largest modulus is only demonstrated here for a 2×2 matrix. However, it works equally well on larger matrices which have real distinct eigenvalues. The advantage of this method over the approach in Section 1 is that iterative schemes like this one are suitable for use on a computer, whereas methods based on the characteristic equation are not.

You will have the opportunity to use the method on larger matrices when you use the computer package described in Section 4. In the meantime, you can see the computer printout we got using the computer package for the 3×3 matrix discussed in the television programme, at the end of this section.

To discover something about the rate of convergence of direct iteration, we look at the algebra of the crude method described initially where scaling was not used.

Let $\mathbf{A}$ be an arbitrary matrix with eigenvalues satisfying the conditions following Procedure 2.4. We start by writing the initial vector $\mathbf{y}_0$ in terms of eigenvectors of $\mathbf{A}$:

$$\mathbf{y}_0 = \mathbf{x}_1 + \mathbf{x}_2.$$

The first iteration gives

$$\begin{aligned}\mathbf{y}_1 &= \mathbf{A}\mathbf{y}_0\\ &= \mathbf{A}(\mathbf{x}_1 + \mathbf{x}_2)\\ &= \lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2.\end{aligned}$$

Similarly, the second iteration gives

$$\begin{aligned}\mathbf{y}_2 &= \mathbf{A}(\lambda_1\mathbf{x}_1 + \lambda_2\mathbf{x}_2)\\ &= \lambda_1(\mathbf{A}\mathbf{x}_1) + \lambda_2(\mathbf{A}\mathbf{x}_2)\\ &= \lambda_1^2\mathbf{x}_1 + \lambda_2^2\mathbf{x}_2.\end{aligned}$$

In general after r iterations we obtain

$$\mathbf{y}_r = \lambda_1^r\mathbf{x}_1 + \lambda_2^r\mathbf{x}_2.$$

If we rewrite this in the form

$$\mathbf{y}_r = \lambda_1^r\left(\mathbf{x}_1 + \left(\frac{\lambda_2}{\lambda_1}\right)^r\mathbf{x}_2\right)$$

and assume that λ_1 is the eigenvalue of largest modulus, then $|\lambda_2/\lambda_1| < 1$, and so $(\lambda_2/\lambda_1)^r$ will become very small as r gets larger. Thus for large enough r,

$$\mathbf{y}_r \simeq \lambda_1^r\mathbf{x}_1.$$

So $\mathbf{y}_r$ is approximately an eigenvector corresponding to the eigenvalue of largest modulus.

The working above shows that if the eigenvalues are well separated, then $|\lambda_2/\lambda_1|$ will be much less than 1, and hence convergence will be rapid.

For instance, in Example 3, you saw that we obtained a good approximation for both the eigenvalue of greatest modulus and its corresponding eigenvector in three iterations. This was because the eigenvalues of $\begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}$ are 1 and 6 so $(1/6)^r$ gets small very rapidly. Unfortunately this is not a fact we know in advance, as the eigenvalues are unknown.

The algebra above gives some indication as to how we would examine convergence, and derive the method for larger matrices, where the geometry discussed in the television programme is inapplicable.

2.5 Inverse iteration

The direct iteration method (or power method) described in the last subsection has the obvious disadvantage that it only finds one eigenvalue (and eigenvector)—the one of largest modulus.

However, in this subsection, we modify this method, and use the modified method to find other selected eigenvalues of a matrix $\mathbf{A}$. First, let us consider the iterative scheme

$$\mathbf{y}_{r+1} = \mathbf{A}^{-1}\mathbf{y}_r \qquad (6)$$

where $\mathbf{A}$ is some non-singular matrix, and $\mathbf{y}_0$ is some arbitrary non-zero vector which is not an eigenvector of $\mathbf{A}$.

From the work in the previous subsection, we know that this scheme will produce a sequence of vectors $\mathbf{y}_1, \mathbf{y}_2, \ldots$ which will tend to the dominant eigenvector of $\mathbf{A}^{-1}$. Further, if we scale this scheme, as in the previous subsection, we will obtain

the eigenvalue of $\mathbf{A}^{-1}$ with *largest* modulus. We know from Subsection 2.1 that the eigenvalues of $\mathbf{A}$ and $\mathbf{A}^{-1}$ are related, so the fact that we can find an eigenvalue μ of $\mathbf{A}^{-1}$, means that we know $1/\mu$ will be an eigenvalue of $\mathbf{A}$. Now, we know that μ is the eigenvalue of $\mathbf{A}^{-1}$ with largest modulus, so $1/\mu$ must be the eigenvalue of $\mathbf{A}$ with *smallest* modulus.

This modified scheme is called inverse iteration and is implemented, including scaling, in the following example.

Example 4

We shall use the scheme described above to find the smallest eigenvalue of

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}.$$

Solution

The inverse of $\mathbf{A}$ is

$$\mathbf{A}^{-1} = \frac{1}{6}\begin{bmatrix} 4 & -2 \\ -3 & 3 \end{bmatrix}.$$

Choosing an initial vector $\mathbf{y}_0 = [1 \quad 1]^T$ and using scaling at each stage, we get

(i) $\mathbf{y}_1' = \frac{1}{6}\begin{bmatrix} 4 & -2 \\ -3 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.333 \\ 0 \end{bmatrix}.$

Dividing by 0.333 we obtain $\mathbf{y}_1 = [1 \quad 0]^T$.

(ii) $\mathbf{y}_2' = \frac{1}{6}\begin{bmatrix} 4 & -2 \\ -3 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.667 \\ -0.5 \end{bmatrix}.$

Dividing by 0.667 we obtain $\mathbf{y}_2 = [1 \quad -0.750]^T$.

(iii) $\mathbf{y}_3' = \frac{1}{6}\begin{bmatrix} 4 & -2 \\ -3 & 3 \end{bmatrix}\begin{bmatrix} 1 \\ -0.75 \end{bmatrix} = \begin{bmatrix} 0.917 \\ -0.875 \end{bmatrix}.$

Dividing by 0.917 we obtain $\mathbf{y}_3 = [1 \quad -0.955]^T$.

Thus, an approximation to the eigenvalue of $\mathbf{A}^{-1}$ with largest modulus is 0.917 and a corresponding eigenvector is $[1 \quad -0.955]^T$. Hence an approximation for the eigenvalue of $\mathbf{A}$ with smallest modulus is $1/0.917 = 1.091$, with a corresponding eigenvector $[1 \quad -0.955]^T$, for we know from Subsection 2.1 that $\mathbf{A}$ and $\mathbf{A}^{-1}$ have the same eigenvectors. The actual eigenvalue is 1, with a corresponding eigenvector $[1 \quad -1]^T$, so this approximation after three iterations is not bad.

We summarize this first version of the inverse iteration procedure below.

Inverse iteration (first version)

Given a non-singular matrix $\mathbf{A}$, and an arbitrary non-zero vector $\mathbf{y}_0$ (not one of the eigenvectors of $\mathbf{A}$)

1. find $\mathbf{y}_{r+1}' = \mathbf{A}^{-1}\mathbf{y}_r$,
2. find $\mathbf{y}_{r+1} = \dfrac{\mathbf{y}_{r+1}'}{\alpha_{r+1}}$,

where α_{r+1} is the element of largest modulus in $\mathbf{y}_{r+1}'$.

Do 1 and 2 for $r = 0, 1, 2, \ldots$.

Then for sufficiently large r, $\dfrac{1}{\alpha_{r+1}}$ will be a good approximation for the eigenvalue of $\mathbf{A}$ with least modulus, and the corresponding eigenvector will be approximately equal to $\mathbf{y}_{r+1}$.

See practical note after next exercise.

Exercise 4

Use the procedure above with $\mathbf{A} = \begin{bmatrix} 7 & 3 \\ 8 & 5 \end{bmatrix}$ and $\mathbf{y}_0 = [1 \quad 1]^T$ to obtain $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_3$. Hence find an approximation to the eigenvalue of $\mathbf{A}$ with least modulus and a corresponding eigenvector.

[*Solution on p. 47*]

Practical Note

For matrices larger than 2×2, the calculation of the inverse matrix is fairly laborious, and the step $\mathbf{y}_{r+1} = \mathbf{A}^{-1}\mathbf{y}_r$ is done by solving the equation $\mathbf{A}\mathbf{y}_{r+1} = \mathbf{y}_r$. However, as the inverse of a 2×2 matrix is very easy to calculate, we advise you to use $\mathbf{A}^{-1}$ when doing these examples by hand.

A more versatile form of inverse iteration

The scheme described above can be generalized to obtain a procedure which will find the eigenvalue of $\mathbf{A}$ which is closest to some chosen value p.

The crude scheme which we use is

$$\mathbf{y}_{r+1} = (\mathbf{A} - p\mathbf{I})^{-1}\mathbf{y}_r. \tag{7}$$

This will, with scaling, certainly determine the eigenvalue μ of $(\mathbf{A} - p\mathbf{I})^{-1}$ which has largest modulus, and this enables us to determine an eigenvalue of $\mathbf{A}$. We know from Subsection 2.1 that if λ is an eigenvalue of $\mathbf{A}$, then $\dfrac{1}{\lambda - p}$ is an eigenvalue of $(\mathbf{A} - p\mathbf{I})^{-1}$. So we can obtain this eigenvalue λ from

$$\mu = \frac{1}{\lambda - p},$$

giving

$$\lambda = \frac{1}{\mu} + p.$$

Now, as μ was the eigenvalue of $(\mathbf{A} - p\mathbf{I})^{-1}$ with largest modulus, $1/\mu$ must be the eigenvalue of $\mathbf{A} - p\mathbf{I}$ with smallest modulus. In other words $1/\mu$ is the eigenvalue of $\mathbf{A} - p\mathbf{I}$ closest to zero and so $\dfrac{1}{\mu} + p$ is the eigenvalue of $\mathbf{A}$ closest to p. The following example demonstrates the use of the Scheme (7), with scaling at each stage.

Example 5

Suppose $\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}$. Find the eigenvalue (and corresponding eigenvector) nearest to 7.

Solution

Putting $p = 7$ gives

$$(\mathbf{A} - p\mathbf{I}) = \begin{bmatrix} -4 & 2 \\ 3 & -3 \end{bmatrix},$$

and so

$$(\mathbf{A} - p\mathbf{I})^{-1} = -\frac{1}{6}\begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}.$$

So we shall implement the Scheme (7), scaling at each stage as before, with an initial vector $\mathbf{y}_0 = [1 \quad 1]^T$.

(i) $$\mathbf{y}_1' = -\frac{1}{6}\begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.833 \\ -1.167 \end{bmatrix}.$$

Dividing by -1.167 we obtain $\mathbf{y}_1 = [0.714 \quad 1]^T$.

(ii) $$\mathbf{y}' = -\frac{1}{6}\begin{bmatrix} 3 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 0.714 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.690 \\ -1.024 \end{bmatrix}.$$

Dividing by -1.024 we obtain $\mathbf{y}_2 = [0.674 \quad 1]^T$.

(iii) $$\mathbf{y}_3' = -\frac{1}{6}\begin{bmatrix} 3 & 2 \\ 2 & 4 \end{bmatrix}\begin{bmatrix} 0.674 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.671 \\ -1.004 \end{bmatrix}.$$

Dividing by -1.004 to obtain $\mathbf{y}_3 = [0.668 \quad 1]^T$.

Thus, an estimate for the eigenvalue of $(\mathbf{A} - p\mathbf{I})^{-1}$ with largest modulus is -1.004, and $[0.668 \quad 1]^T$ is the corresponding eigenvector. Hence, an estimate for the eigenvalue of $\mathbf{A}$ nearest to 7 is $\dfrac{1}{-1.004} + 7 = 6.004$. The corresponding eigenvector is $[0.668 \quad 1]^T$ for we know that $\mathbf{A}$ and $(\mathbf{A} - p\mathbf{I})^{-1}$ have the same eigenvectors. Again, these turn out to be good estimates for the actual eigenvalue 6 and eigenvector $[2/3 \quad 1]^T$.

This scheme, as formally set out below, is the one we shall refer to as **inverse iteration**.

Procedure 2.5: Inverse iteration

To find the eigenvalue closest to p and the corresponding eigenvector of a given matrix $\mathbf{A}$, start with an arbitrary non-zero vector $\mathbf{y}_0$ (not one of the eigenvectors of $\mathbf{A}$) and then

1. find $\mathbf{y}_{r+1}' = (\mathbf{A} - p\mathbf{I})^{-1}\mathbf{y}_r$,
2. find $\mathbf{y}_{r+1} = \dfrac{\mathbf{y}_{r+1}'}{\alpha_{r+1}}$,

where α_{r+1} is the element of largest modulus in $\mathbf{y}_{r+1}'$.
Do 1 and 2 for $r = 0, 1, 2, \ldots$.

Then for sufficiently large r, $\dfrac{1}{\alpha_{r+1}} + p$ will be a good approximation for the eigenvalue of $\mathbf{A}$ nearest to p, and $\mathbf{y}_{r+1}$ will be a good approximation for the corresponding eigenvector.

See practical note on page 23.

This result in fact covers the case of the previous boxed result as, by putting $p = 0$, we get the eigenvalue nearest to 0, that is, the eigenvalue of $\mathbf{A}$ with smallest modulus.

Exercise 5

Find $\mathbf{y}_1$, $\mathbf{y}_2$ and $\mathbf{y}_3$ using the scheme above, where $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$, $p = 2$, and using a starting vector $\mathbf{y}_0 = [1 \quad 0]^T$. Hence, estimate the eigenvalue of $\mathbf{A}$ nearest to its corresponding eigenvector.

[*Solution on p. 48*]

A note about choosing p for a 3×3 matrix

For a 3×3 matrix the largest and smallest eigenvalues can be found by iterating with $\mathbf{A}$ and $\mathbf{A}^{-1}$ respectively. The remaining eigenvalue can be found by iterating with $(\mathbf{A} - p\mathbf{I})^{-1}$ provided a suitable value can be chosen for p.

In the television programme the three eigenvalues of the matrix

$$\begin{bmatrix} 6.7 & 3 & 2.1 \\ 4.5 & -5.1 & 0.5 \\ 2.3 & -5.6 & 9.5 \end{bmatrix}$$

are found, using the computer package EIGSOL described in Section 4. The eigenvalues of largest and smallest modulus were found to be approximately 9.91 and -6.24 respectively.

Figure 6

Having found these two eigenvalues the choice of a suitable p can be cut down to two regions (see Figure 6). The remaining eigenvalue cannot lie between -6.24 and 6.24—for otherwise it (and not 6.24) would be the eigenvalue of smallest modulus. Nor can it be larger than 9.91 or smaller than -9.91, for then it (and not 9.91) would be the eigenvalue of largest modulus. So the remaining eigenvalue must either lie between 6.24 and 9.91, or between -6.24 and -9.91. So try $p = (6.24 + 9.91)/2 \simeq 8$, or if this does not work try $p = -8$. One of these values of p will give you the remaining eigenvalue.

In general, if λ_1 and λ_2 are the eigenvalues of largest and smallest modulus respectively for some 3×3 matrix, then choosing p to be either $\dfrac{|\lambda_1| + |\lambda_2|}{2}$ or $-\left(\dfrac{|\lambda_1| + |\lambda_2|}{2}\right)$ is bound to produce the remaining eigenvalue (see Figure 7).

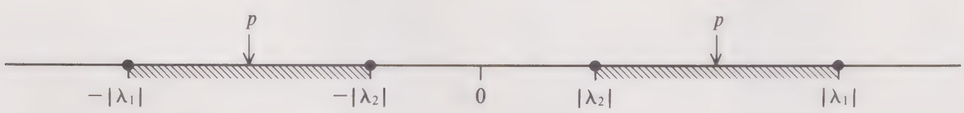

Figure 7

Finally, here is the computer output we got when finding the three eigenvalues (and eigenvectors) of the matrix used in the television programme.

(i) DIRECT ITERATION

n	Yn			ALPHA
1	1	−0.00848	0.52542	11.8000
6	1	0.38486	0.97957	9.33112
11	0.98047	0.32369	1	9.87006
16	0.95930	0.32176	1	9.91254
21	0.95542	0.31986	1	9.90566
26	0.95439	0.31955	1	9.90570
31	0.95415	0.31947	1	9.90555
36	0.95410	0.31944	1	9.90553
41	0.95408	0.31944	1	9.90552
45	0.95408	0.31944	1	9.90552

Eigenvalue of largest modulus is 9.90552
Found in 45 iterations

Corresponding eigenvector

0.95408
0.31944
1

(ii) INVERSE ITERATION (WITH p = 0)

n	Yn			ALPHA
1	1	−0.31642	0.22687	0.16059
11	−0.41539	1	0.42106	−0.14796
21	−0.31521	1	0.40413	−0.15832
31	−0.29990	1	0.40010	−0.16001
41	−0.29728	1	0.39940	−0.16031
51	−0.29683	1	0.39928	−0.16036
61	−0.29675	1	0.39926	−0.16037
71	−0.29673	1	0.39925	−0.16037
81	−0.29673	1	0.39925	−0.16037
86	−0.29673	1	0.39925	−0.16037

The eigenvalue of maximum modulus of the inverse of
(A − 0.00 *I) is −0.16037
Hence the eigenvalue nearest to 0.00 is −6.23567

Found in 86 iterations

Corresponding eigenvector

−0.29673
1
0.39925

(iii) INVERSE ITERATION (WITH p = 8)

n	Yn			ALPHA
1	−0.50743	−0.28218	1	0.52271
2	1	0.35244	0.02694	2.72664
3	1	0.35402	−0.22438	−1.41021
4	1	0.35261	−0.13750	−1.88353
5	1	0.35298	−0.16222	−1.71900
6	1	0.35286	−0.15472	−1.76578
7	1	0.35290	−0.15696	−1.75159
8	1	0.35289	−0.15629	−1.75582
9	1	0.35289	−0.15649	−1.75455
10	1	0.35289	−0.15643	−1.75493
11	1	0.35289	−0.15645	−1.75482
12	1	0.35289	−0.15644	−1.75485
13	1	0.35289	−0.15644	−1.75484
14	1	0.35289	−0.15644	−1.75484

The eigenvalue of maximum modulus of the inverse of
(A − 8.00 *I) is −1.75484
Hence the eigenvalue nearest to 8.00 is 7.43015

Found in 14 iterations

Corresponding eigenvector

1
0.35289
−0.15644

Summary of Section 2

1. If **A** is a matrix with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ then:

(i) $\mathbf{A} + q\mathbf{I}$ has eigenvalues $\lambda_1 + q, \lambda_2 + q, \ldots, \lambda_n + q$;

(ii) $\mathbf{A}^{-1}$ has eigenvalues $\frac{1}{\lambda_1}, \frac{1}{\lambda_2}, \ldots, \frac{1}{\lambda_n}$ so long as **A** is non-singular;

(iii) $(\mathbf{A} - p\mathbf{I})^{-1}$ has eigenvalues $\dfrac{1}{\lambda_1 - p}, \dfrac{1}{\lambda_2 - p}, \ldots, \dfrac{1}{\lambda_n - p}$ so long as $\mathbf{A} - p\mathbf{I}$ is non-singular;

(iv) $\mathbf{A}$, $\mathbf{A} + q\mathbf{I}$, $\mathbf{A}^{-1}$ and $(\mathbf{A} - p\mathbf{I})^{-1}$ all have the same eigenvectors.

2. If a 2×2 matrix $\mathbf{A}$ has two real eigenvectors, then there are special directions $\mathbf{x}$ such that the vector $\mathbf{Ax}$ is in the same direction as $\mathbf{x}$. These special directions are given by the eigenvectors of $\mathbf{A}$. The factor by which $\mathbf{A}$ stretches the original vector $\mathbf{x}$ is given by the corresponding eigenvalues of $\mathbf{A}$.

3. The procedures summarized below determine single eigenvalues of a matrix $\mathbf{A}$. In all cases we start with an arbitrary non-zero vector $\mathbf{y}_0$ (not an eigenvector). α_{r+1} is the element of largest modulus in $\mathbf{y}'_{r+1}$.

Procedure	Iterative scheme to be used for $r = 0, 1, 2, \ldots$	Result for large r
Direct iteration	Form $\mathbf{y}'_{r+1} = \mathbf{A}\mathbf{y}_r$ and set $\mathbf{y}_{r+1} = \dfrac{\mathbf{y}'_{r+1}}{\alpha_{r+1}}$	α_{r+1} approximates the eigenvalue of largest modulus. $\mathbf{y}_{r+1}$ approximates the corresponding eigenvector.
Inverse iteration	Form $\mathbf{y}'_{r+1} = \mathbf{A}^{-1}\mathbf{y}_r$ and set $\mathbf{y}_{r+1} = \dfrac{\mathbf{y}'_{r+1}}{\alpha_{r+1}}$	$\dfrac{1}{\alpha_{r+1}}$ approximates the eigenvalue of smallest modulus. $\mathbf{y}_{r+1}$ approximates the corresponding eigenvector.
Modified inverse iteration	Form $\mathbf{y}'_{r+1} = (\mathbf{A} - p\mathbf{I})^{-1}\mathbf{y}_r$ and set $\mathbf{y}_{r+1} = \dfrac{\mathbf{y}'_{r+1}}{\alpha_{r+1}}$	$\dfrac{1}{\alpha_{r+1}} + p$ approximates the nearest eigenvalue to a chosen value p. $\mathbf{y}_{r+1}$ approximates the corresponding eigenvector.

Note that the second iterative scheme is a special case of the third scheme with $p = 0$.

The above schemes will only work if the eigenvalue we are looking for is real and distinct from all the others.

4. For a 3×3 matrix, if the eigenvalues of largest and smallest modulus are λ_1 and λ_2, then to find the remaining eigenvalue, choose $p = \pm\left(\dfrac{|\lambda_1| + |\lambda_2|}{2}\right)$.

End of section exercise

Exercise 6

For the matrix

$$\mathbf{A} = \begin{bmatrix} -39 & 40 \\ -20 & 21 \end{bmatrix}$$

use direct and inverse iteration with initial vector $\mathbf{y}_0 = [1 \quad 0]^T$ to find:

(i) the eigenvalue of $\mathbf{A}$ with largest modulus;

(ii) the eigenvalue of $\mathbf{A}$ nearest to 2.

In each case, stop the method when each element in two successive approximations $\mathbf{y}_{r+1}$ and $\mathbf{y}_r$ agree to two significant figures.

[*Solution on p. 48*]

3 Decomposition methods for finding all the eigenvalues

3.0 Introduction

In Section 2, you saw the development of a class of iterative methods, which can be used to find selected eigenvalues (and associated eigenvectors) of a given matrix **A**. Often, however, in practical problems, all the eigenvalues are required. Although it would be possible in theory to find these eigenvalues by repeated use of inverse iteration with suitable values of p, it is in fact more practical to use a method which is specifically designed to find all the eigenvalues simultaneously. Most methods of this type currently in use depend on suitable 'decompositions' in which the given matrix **A** is expressed in the form **PQ** for suitable matrices **P** and **Q**. It is for this reason that Subsection 3.1 concentrates on matrix decomposition. Subsection 3.2 deals with some algebraic results which will help you understand the method described in Subsection 3.3. This is a method which can be used to find all the eigenvalues of a matrix simultaneously.

The exercises that you will be asked to do in this section are mostly on 2×2 matrices. This is simply to cut down the amount of hard computation you are required to do. Rest assured that the processes described will work equally well for larger matrices, and you will have the opportunity to do some of these using the computer package in Section 4.

3.1 Matrix decomposition

Decomposition of matrices is analogous to factorization of numbers. For instance much in the same way that 12 has a factorization $12 = 3 \times 4$ you can check that the matrix $\begin{bmatrix} 1 & 1 \\ 1 & 6 \end{bmatrix}$ has a decomposition

$$\begin{bmatrix} 1 & 1 \\ 1 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & \sqrt{5} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & \sqrt{5} \end{bmatrix}. \tag{1}$$

Now, some numbers have more than one factorization. For instance $12 = 3 \times 4$, or $12 = 6 \times 2$. Similarly it is possible for a matrix to have more than one decomposition. For instance, you can check that the matrix above has another decomposition

$$\begin{bmatrix} 1 & 1 \\ 1 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 0 & 5 \end{bmatrix}. \tag{2}$$

In this subsection, we discuss just one possible decomposition that a given matrix may have. In certain circumstances, a square matrix **A** can be decomposed as the product

$$\mathbf{A} = \mathbf{LU},$$

where **L** is a lower triangular matrix, which is chosen to have ones along its main diagonal, and **U** is an ordinary upper triangular matrix. This is often referred to as an **LU decomposition**, and an example is given below.

Example 1

$$\underset{\mathbf{A}}{\begin{bmatrix} 3 & 1 & 1 \\ -3 & 5 & 1 \\ 6 & -4 & 1 \end{bmatrix}} = \underset{\mathbf{L}}{\begin{bmatrix} 1 & 0 & 0 \\ -1 & 1 & 0 \\ 2 & -1 & 1 \end{bmatrix}} \underset{\mathbf{U}}{\begin{bmatrix} 3 & 1 & 1 \\ 0 & 6 & 2 \\ 0 & 0 & 1 \end{bmatrix}}$$

Another example of an **LU** decomposition is the 2×2 decomposition (2) above.

A simple approach to finding an **LU** decomposition is just to assume it exists! The following example demonstrates this technique.

Example 2

To find the **LU** decomposition of

$$\mathbf{A} = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix},$$

assume that

$$\underset{\mathbf{A}}{\begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}} = \underset{\mathbf{L}}{\begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}} \underset{\mathbf{U}}{\begin{bmatrix} b & c \\ 0 & d \end{bmatrix}}.$$

To find the unknown elements a, b, c and d we shall just multiply **L** and **U** together, to obtain

$$\begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} b & c \\ ab & ac + d \end{bmatrix}. \tag{3}$$

Then equating these two matrices, element by element, we get from the first row

$$b = 5, \qquad c = 3,$$

and from the second row

$$ab = 3, \qquad ac + d = 5.$$

Hence, $a = \frac{3}{5}$ and $d = \frac{16}{5}$, and the required decomposition is

$$\begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{3}{5} & 1 \end{bmatrix} \begin{bmatrix} 5 & 3 \\ 0 & \frac{16}{5} \end{bmatrix}.$$

Exercise 1

Find the **LU** decomposition of the matrix

$$\mathbf{A} = \begin{bmatrix} 6 & 5 \\ 4 & 5 \end{bmatrix}.$$

[*Solution on p. 48*]

Not all matrices have an **LU** decomposition. Given a particular matrix, which has no **LU** decomposition, the arithmetic will break down when we attempt to do the decomposition, as demonstrated in the following example.

Example 3

Show that $\begin{bmatrix} 0 & 1 \\ -3 & 4 \end{bmatrix}$ does not have an **LU** decomposition.

Solution

If there were an **LU** decomposition we would have

$$\begin{bmatrix} 0 & 1 \\ -3 & 4 \end{bmatrix} = \begin{bmatrix} b & c \\ ab & ac + d \end{bmatrix} \qquad \text{(from (3) above).}$$

Equating coefficients, from the top row we obtain $b = 0$ and $c = 1$. Equating coefficients in the second row, we obtain $ab = -3$ and $ac + d = 5$. Now, it is not possible for ab to be -3, as b is zero. Thus, the arithmetic process breaks down, and we know that no such decomposition exists.

In general, a 2×2 matrix has an **LU** decomposition

$$\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \dfrac{a_{21}}{a_{11}} & 1 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} \\ 0 & \dfrac{1}{a_{11}} \det \mathbf{A} \end{bmatrix}, \tag{4}$$

so long as a_{11} is not zero.

Remember: $\det \mathbf{A} = a_{11}a_{22} - a_{12}a_{21}$.

You can check the result above by multiplying out the right-hand side. This result has been boxed and labelled (4) so that we can use it in later examples.

To find an **LU** decomposition of an $n \times n$ matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix} = \begin{bmatrix} 1 & & & \\ l_{21} & 1 & & \mathbf{0} \\ \vdots & \vdots & \ddots & \\ l_{n1} & l_{n2} & \cdots & 1 \end{bmatrix} \begin{bmatrix} u_{11} & \cdots & u_{1n} \\ & \ddots & \vdots \\ \mathbf{0} & & u_{nn} \end{bmatrix}$$

Notation: all the matrix elements replaced by $\mathbf{0}$ are zero.

is in principle no more difficult than for a 2×2 matrix. It just involves more arithmetic. There are many computer packages that will do this for you.

Exercise 2

Find the **LU** decomposition of the matrix

$$\mathbf{A} = \begin{bmatrix} 1 & 2 & 3 \\ 1 & 6 & 4 \\ 2 & 1 & 4 \end{bmatrix}.$$

[*Solution on p. 48*]

3.2 Some algebraic results

To gain some insight as to why the process described in the next subsection works, we shall need some properties of eigenvalues. For easy reference, these will be referred to as theorems.

The first theorem concerns the eigenvalues of the matrices **AB** and **BA**. The theorem asserts that these two matrices have the same eigenvalues, as the following example illustrates.

Example 4

Let $\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix}$ and $\mathbf{B} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix}$. Then

$$\mathbf{AB} = \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ 2 & 3 \end{bmatrix}.$$

So the eigenvalues of **AB** can be found by solving

$$\begin{vmatrix} 2-\lambda & 1 \\ 2 & 3-\lambda \end{vmatrix} = 0.$$

That is,

$$(2-\lambda)(3-\lambda) - 2 = 0 \qquad (5)$$

or $\quad \lambda^2 - 5\lambda + 4 = 0$

i.e. $\quad (\lambda - 4)(\lambda - 1) = 0.$

Thus the eigenvalues of **AB** are 4 and 1.

Now consider **BA**.

$$\mathbf{BA} = \begin{bmatrix} 2 & 1 \\ 0 & 2 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3 & 1 \\ 2 & 2 \end{bmatrix}$$

and so the eigenvalues of **BA** can be found by solving

$$\begin{vmatrix} 3-\lambda & 1 \\ 2 & 2-\lambda \end{vmatrix} = 0.$$

This gives $(3-\lambda)(2-\lambda) - 2 = 0$, which is the same as Equation (5). So the eigenvalues of **AB** are also 4 and 1.

Hence, in this case, **AB** and **BA** have the same eigenvalues. In general, for any matrices **A** and **B**, the eigenvalues of **AB** and **BA** are the same, but their eigenvectors are usually different.

Exercise 3

Show that $\mathbf{AB}$ and $\mathbf{BA}$ have the same eigenvalues if

$$\mathbf{A} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} 1 & 4 \\ 0 & -10 \end{bmatrix}.$$

[*Solution on p. 49*]

The general result can be obtained from the definition of an eigenvalue. Let $\mathbf{A}$ and $\mathbf{B}$ be any square matrices. If λ is an eigenvalue of $\mathbf{AB}$ with corresponding eigenvector $\mathbf{x}$ then

$$\mathbf{ABx} = \lambda\mathbf{x}.$$

If we multiply this equation on the left by $\mathbf{B}$, we get

$$\mathbf{BABx} = \lambda\mathbf{Bx}.$$

Now, $\mathbf{Bx}$ is just a column vector. So letting $\mathbf{y} = \mathbf{Bx}$, we obtain the equation

$$\mathbf{BAy} = \lambda\mathbf{y}.$$

From this equation, we can see that λ is also an eigenvalue of $\mathbf{BA}$ with a corresponding eigenvector $\mathbf{y} = \mathbf{Bx}$. A similar argument, starting with $\mathbf{BA}$, shows that if λ is an eigenvalue of $\mathbf{BA}$ then λ is also an eigenvalue of $\mathbf{AB}$. Thus we have derived the following theorem.

Theorem 1

For any square matrices $\mathbf{A}$ and $\mathbf{B}$.

(i) $\mathbf{AB}$ and $\mathbf{BA}$ have the same eigenvalues;

(ii) If $\mathbf{x}$ is an eigenvector of $\mathbf{AB}$ then $\mathbf{Bx}$ is an eigenvector of $\mathbf{BA}$.

This result sometimes appears in other forms. For instance, suppose $\mathbf{P}$ and $\mathbf{X}$ are square matrices and that $\mathbf{P}$ is non-singular. If we replace $\mathbf{A}$ by $\mathbf{XP}$ and $\mathbf{B}$ by $\mathbf{P}^{-1}$ in the result above, then $\mathbf{AB} = (\mathbf{XP})\mathbf{P}^{-1} = \mathbf{X}$ and $\mathbf{BA} = \mathbf{P}^{-1}\mathbf{XP}$. Thus we can derive the following form of the result above.

Theorem 2

For any square matrices $\mathbf{P}$ and $\mathbf{X}$ with $\mathbf{P}$ non-singular:

(i) $\mathbf{X}$ and $\mathbf{P}^{-1}\mathbf{XP}$ have the same eigenvalues;

(ii) If $\mathbf{x}$ is an eigenvector of $\mathbf{X}$ then $\mathbf{P}^{-1}\mathbf{x}$ is an eigenvector of $\mathbf{P}^{-1}\mathbf{XP}$.

This result will be used in *Unit 22*.

Some more results which we shall need are much easier to show.

Theorem 3

The eigenvalues of the diagonal matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & & 0 \\ & \ddots & \\ 0 & & a_{nn} \end{bmatrix}$$

are $a_{11}, a_{22}, \ldots, a_{nn}$.

This follows immediately from finding the eigenvalues from $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$.

$$\det(\mathbf{A} - \lambda\mathbf{I}) = \begin{vmatrix} a_{11} - \lambda & & 0 \\ & \ddots & \\ 0 & & a_{nn} - \lambda \end{vmatrix}$$
$$= (a_{11} - \lambda)(a_{22} - \lambda)\ldots(a_{nn} - \lambda).$$

Hence we find that the eigenvalues of **A** are $a_{11}, a_{22}, \ldots, a_{nn}$. In a very similar fashion we can derive the following more general result.

Theorem 4

The eigenvalues of the upper triangular matrix

$$\mathbf{A} = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} \end{bmatrix}$$

are $a_{11}, a_{22}, \ldots, a_{nn}$.

This again follows from putting $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$. Now the matrix

$$\mathbf{A} - \lambda\mathbf{I} = \begin{bmatrix} a_{11} - \lambda & \cdots & a_{1n} \\ & \ddots & \vdots \\ 0 & & a_{nn} - \lambda \end{bmatrix}$$

is upper triangular and so its determinant is equal to the product of the diagonal elements. That is

$$\det(\mathbf{A} - \lambda\mathbf{I}) = (a_{11} - \lambda)(a_{22} - \lambda)\ldots(a_{nn} - \lambda).$$

Determinants of upper triangular matrices were discussed in *Unit 20*, Subsection 5.4.

Hence, once again we obtain the result that the eigenvalues of an upper triangular matrix are its diagonal elements $a_{11}, a_{22}, \ldots, a_{nn}$.

Exercise 4

State the eigenvalues of the following matrices

(i) $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ (ii) $\begin{bmatrix} 7 & 2 & 1 \\ 0 & 8 & 5 \\ 0 & 0 & 9 \end{bmatrix}$

[*Solution on p. 49*]

The last theorem we state in this subsection is a result you will require for *Unit 22*.

Theorem 5

Suppose that a square matrix **A** has n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Let **P** be a matrix whose columns are corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ of **A**. Then

$$\mathbf{P}^{-1}\mathbf{A}\mathbf{P} = \begin{bmatrix} \lambda_1 & & & 0 \\ & \lambda_2 & & \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}.$$

If the eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$ are not all distinct, the result still holds provided we can choose a corresponding set of eigenvectors which are linearly independent.

This result is illustrated in the following example.

Example 5

Using the results of Example 5 in Subsection 1.2 we know that $\mathbf{A} = \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}$ has eigenvalues 3 and 7, with corresponding eigenvectors $[-1 \quad 1]^T$ and $[1 \quad 1]^T$. So, we know that

$$\begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} -1 \\ 1 \end{bmatrix} = 3\begin{bmatrix} -1 \\ 1 \end{bmatrix}, \text{ and } \begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = 7\begin{bmatrix} 1 \\ 1 \end{bmatrix}.$$

We can put these two equations as a single matrix equation in the following way:

$$\begin{bmatrix} 5 & 2 \\ 2 & 5 \end{bmatrix}\begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix} = \begin{bmatrix} 3\boxed{\begin{matrix} -1 \\ 1 \end{matrix}} & 7\boxed{\begin{matrix} 1 \\ 1 \end{matrix}} \end{bmatrix}$$
$$= \begin{bmatrix} -1 & 1 \\ 1 & 1 \end{bmatrix}\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}.$$

So letting $\mathbf{P}$ be the matrix whose columns are the eigenvectors of $\mathbf{A}$, we have that

$$\mathbf{AP} = \mathbf{P}\begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix},$$

or

$$\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} 3 & 0 \\ 0 & 7 \end{bmatrix}.$$

This matrix just has the eigenvalues of $\mathbf{A}$ on the main diagonal.

The general result of Theorem 5 will not be proved in this unit, but the proof is very similar to the example above.

Exercise 5

Given that $\mathbf{A} = \begin{bmatrix} 2 & 3 \\ 3 & 2 \end{bmatrix}$, find matrices $\mathbf{P}$ and $\mathbf{D}$ such that

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D},$$

where $\mathbf{D}$ is a diagonal matrix consisting of the eigenvalues of $\mathbf{A}$.

[*Solution on p. 49*]

3.3 The LR method

The LU decomposition of a matrix is used in a process called the LR algorithm for finding eigenvalues. First, suppose we have a matrix $\mathbf{A}_0$, and we find its LU decomposition.

$$\mathbf{A}_0 = \mathbf{L}_0\mathbf{U}_0.$$

If we then form a new matrix

$$\mathbf{A}_1 = \mathbf{U}_0\mathbf{L}_0,$$

we know that

(i) In general $\mathbf{A}_0$ and $\mathbf{A}_1$ will be different matrices, as $\mathbf{L}_0\mathbf{U}_0 \neq \mathbf{U}_0\mathbf{L}_0$.

(ii) From Theorem 1, $\mathbf{A}_0$ and $\mathbf{A}_1$ have the same eigenvalues.

The next example illustrates how repeated use of this process leads to a method for approximating the eigenvalues of a matrix.

Example 6

Suppose we want to find the eigenvalues of the matrix

$$\mathbf{A}_0 = \begin{bmatrix} 5 & 3 \\ 3 & 5 \end{bmatrix}.$$

We start by finding the decomposition $\mathbf{A}_0 = \mathbf{L}_0\mathbf{U}_0$:

$$\mathbf{A}_0 = \begin{bmatrix} 1 & 0 \\ 0.6 & 1 \end{bmatrix}\begin{bmatrix} 5 & 3 \\ 0 & 3.2 \end{bmatrix} \quad \text{(from Equation (4) of Subsection 3.2)}$$

and then form a new matrix $\mathbf{A}_1 = \mathbf{U}_0\mathbf{L}_0$:

$$\mathbf{A}_1 = \begin{bmatrix} 5 & 3 \\ 0 & 3.2 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0.6 & 1 \end{bmatrix}$$
$$= \begin{bmatrix} 6.8 & 3 \\ \mathbf{1.92} & 3.2 \end{bmatrix}.$$

We know from Theorem 1 that $\mathbf{A}_0$ and $\mathbf{A}_1$ have the same eigenvalues. By repeating this process we can form a sequence of matrices $\mathbf{A}_2, \mathbf{A}_3, \ldots$, each of which has the same eigenvalues as $\mathbf{A}_0$. Let us see what happens when we do this.

All calculations here and in the rest of this unit are worked to the full accuracy of the computer. For convenience, the numbers have been truncated in the text.

The **LU** decomposition of $\mathbf{A}_1$ is

$$\mathbf{A}_1 = \mathbf{L}_1\mathbf{U}_1 = \begin{bmatrix} 1 & 0 \\ 0.282 & 1 \end{bmatrix}\begin{bmatrix} 6.8 & 3 \\ 0 & 2.353 \end{bmatrix}$$

(from Equation (4) of Subsection 3.2),

and so the next matrix $\mathbf{A}_2$ is given by

$$\mathbf{A}_2 = \mathbf{U}_2\mathbf{L}_2 = \begin{bmatrix} 6.8 & 3 \\ 0 & 2.353 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0.282 & 1 \end{bmatrix} = \begin{bmatrix} 7.647 & 3 \\ \mathbf{0.664} & 2.353 \end{bmatrix}.$$

The next iteration produces the **LU** decomposition

$$\mathbf{A}_2 = \mathbf{L}_2\mathbf{U}_2 = \begin{bmatrix} 1 & 0 \\ 0.0869 & 1 \end{bmatrix}\begin{bmatrix} 7.647 & 3 \\ 0 & 2.092 \end{bmatrix},$$

and so $\mathbf{A}_3$ is given by

$$\mathbf{A}_3 = \mathbf{U}_2\mathbf{L}_2 = \begin{bmatrix} 7.647 & 3 \\ 0 & 2.092 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0.0869 & 1 \end{bmatrix} = \begin{bmatrix} 7.908 & 3 \\ \mathbf{0.182} & 2.092 \end{bmatrix}.$$

It is worth noting that in this sequence of matrices $\mathbf{A}_1, \mathbf{A}_2, \mathbf{A}_3, \ldots$, which we are forming, the bottom left-hand element (set in bold type) appears to be getting smaller. If this continues to happen in each subsequent iteration, then sooner or later, some matrix $\mathbf{A}_r$ which we form will virtually be in upper triangular form. If this happens, then we know from Theorem 4 that the eigenvalues of $\mathbf{A}_r$, and hence of $\mathbf{A}_0$, will be the values on the diagonal. Performing one more iteration, we obtain the decomposition

$$\mathbf{A}_3 = \mathbf{L}_3\mathbf{U}_3 = \begin{bmatrix} 1 & 0 \\ 0.023 & 1 \end{bmatrix}\begin{bmatrix} 7.908 & 3 \\ 0 & 2.023 \end{bmatrix},$$

and so $\mathbf{A}_4$ is given by

$$\mathbf{A}_4 = \mathbf{U}_3\mathbf{L}_3 = \begin{bmatrix} 7.908 & 3 \\ 0 & 2.023 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0.023 & 1 \end{bmatrix} = \begin{bmatrix} 7.977 & 3 \\ \mathbf{0.047} & 2.023 \end{bmatrix}.$$

Again you can see the bottom left-hand element getting smaller. If this element is taken to be approximately zero, then the matrix is approximately diagonal, and so by Theorem 4 the eigenvalues are approximately equal to the elements on the main diagonal.

Thus we obtain approximate eigenvalues of 7.977 and 2.023 for the sequences $\mathbf{A}_r$ of matrices which we have formed. In particular, these are approximate eigenvalues of the matrix $\mathbf{A}_0$ which we started with. In fact, the eigenvalues of $\mathbf{A}_0$ are 8 and 2, and clearly, had we continued with the process described above, we would have got better approximations to these values.

The method described above can be expressed as an iterative procedure. Starting with a square matrix $\mathbf{A}_0$ we form a sequence of matrices $\mathbf{A}_r$, $r = 0, 1, 2, \ldots$ At each stage we find the LU decomposition

$$\mathbf{A}_r = \mathbf{L}_r\mathbf{U}_r$$

and use this to form the next matrix

$$\mathbf{A}_{r+1} = \mathbf{U}_r\mathbf{L}_r.$$

This matrix has the same eigenvalues as $\mathbf{A}_r$.

It can be shown that under certain conditions, as r increases, $\mathbf{A}_r$ approaches an upper triangular matrix, with its eigenvalues (and hence those of $\mathbf{A}_0$) in descending order of modulus on the main diagonal. The conditions for which the process converges are very complicated, and the proof is beyond the scope of this unit. For instance, you know that in some cases an LU decomposition does not exist, and certainly in those cases we cannot carry out the procedure.

My advice to you is to first try the procedure and see if it works. If not, try the procedure on a matrix of the form $\mathbf{A} + q\mathbf{I}$, as you know from Subsection 2.1 that this will produce eigenvalues $\lambda_i + q$, where λ_i are the eigenvalues you want.

This method is usually known as the **LR method** or **algorithm** (rather than the LU method) for this is the way it was referred to when first introduced by Rutishauser in 1958. In his original paper, he called the upper triangular matrix $\mathbf{R}$, rather than the $\mathbf{U}$ adopted in this text.

Procedure 3.3: The LR method

Given a matrix $\mathbf{A}_0$, form a sequence of matrices $\mathbf{A}_r$ $(r = 0, 1, 2, \ldots)$ using the following procedure:

1. Find the decomposition

$$\mathbf{A}_r = \mathbf{L}_r\mathbf{U}_r$$

where $\mathbf{L}_r$ is a lower triangular matrix with ones on its main diagonal, and $\mathbf{U}_r$ is an upper triangular matrix.

2. Form the new matrix

$$\mathbf{A}_{r+1} = \mathbf{U}_r\mathbf{L}_r.$$

Do 1 and 2 for $r = 0, 1, 2, \ldots$

Then, for suitable matrices $\mathbf{A}_0$, and large enough r, $\mathbf{A}_{r+1}$ converges to an upper triangular matrix, with the eigenvalues in order of descending modulus on the diagonal.

Comment

If the LR method does not work on $\mathbf{A}_0$, do the same process on $\mathbf{A}_0 + q\mathbf{I}$ for some suitable choice of q. In this event, the eigenvalues you find will be $\lambda_i + q$ where λ_i are the eigenvalues of $\mathbf{A}_0$.

Exercise 6

Use three iterations of the method described above to find an approximation for the eigenvalues of the matrix

$$\mathbf{A}_0 = \begin{bmatrix} 6 & 5 \\ 4 & 5 \end{bmatrix}.$$

[*Solution on p. 49*]

It should be stressed that this method is for use on a computer and is not meant for hand calculation. It is for this reason that there are not many exercises in this subsection. You will be given the opportunity to see the LR method in practice when you use the computer package EIGSOL, which is described in Section 4.

This section is only an introduction to a number of methods, called **decomposition methods**, which are currently used for finding eigenvalues. All these methods decompose a matrix into two matrices $\mathbf{P}$ and $\mathbf{Q}$, and form a new matrix $\mathbf{QP}$. The particular example of an LU decomposition was chosen, as it is one of the easiest to describe. In the crude form described in this unit, it suffers from slow convergence and hence from accumulation of rounding error. The way in which the method is amended to speed up convergence is beyond the scope of this unit. If you want more information on the numerical methods described in these sections, consult *The Algebraic Eigenvalue Problem* by J. H. Wilkinson (Oxford University Press, 1965) which is probably the most comprehensive book to date on this topic.

Summary of Section 3

1. If two matrices can be found such that

$$\mathbf{A} = \mathbf{PQ}$$

All the matrices must be square.

we say that $\mathbf{A}$ can be **decomposed**, and the whole process is referred to as a **decomposition**.

2. The **LU decomposition** of a matrix $\mathbf{A}$ is

$$\mathbf{A} = \mathbf{LU}$$

where $\mathbf{L}$ is a lower triangular matrix with ones down the main diagonal, and $\mathbf{U}$ an upper triangular matrix. Not all square matrices have LU decompositions.

3. Some important matrix results are:

(i) $\mathbf{AB}$ and $\mathbf{BA}$ have the same eigenvalues.

(ii) $\mathbf{A}$ and $\mathbf{P}^{-1}\mathbf{AP}$ have the same eigenvalues.

(iii) The eigenvalues of a diagonal matrix are the elements on the main diagonal.

(iv) The eigenvalues of an upper triangular matrix are the elements on the main diagonal.

(v) Let $\mathbf{A}$ be a square matrix, with eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, and $\mathbf{P}$ the matrix whose columns are the corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$ of $\mathbf{A}$. Then

$$\mathbf{P}^{-1}\mathbf{AP} = \begin{bmatrix} \lambda_1 & & & \\ & \lambda_2 & & 0 \\ & & \ddots & \\ 0 & & & \lambda_n \end{bmatrix}$$

4. The **LR method** is a method for finding all the eigenvalues of certain matrices. The method is described in Procedure 3.4. It is an example of a **decomposition method**.

End of section exercise

Exercise 7

Do three iterations of the LR algorithm on the matrix

$$\mathbf{A}_0 = \begin{bmatrix} 7 & 5 \\ 3 & 5 \end{bmatrix}.$$

Hence make an approximation of the eigenvalues of $\mathbf{A}_0$.

[*Solution on p. 49*]

4 The computer package EIGSOL

4.0 Introduction

This section describes the computer package EIGSOL.

EIGSOL is designed to find eigenvalues and eigenvectors of an $n \times n$ matrix. Again, all the remarks about computer packages in Section 6 of *Unit 1* are still valid.

As with other computer packages described in this course, the options are split into five main categories. However, before giving a full list of options, we shall look at a worked example.

4.1 A worked example

In the television programme we used the computer package EIGSOL to find the eigenvalue of largest modulus, and the corresponding eigenvector, of the matrix

$$\begin{bmatrix} 6.7 & 3 & 2.1 \\ 4.5 & -5.1 & 0.5 \\ 2.3 & -5.6 & 9.5 \end{bmatrix}.$$

To do this, only a subset of the options described in Subsection 4.3 is required. The following computer dialogue illustrates the options needed for this particular problem. The information you have to type in is underlined. The rest of the information is supplied by the computer.

When you have logged on, and obtained the computer package EIGSOL, the computer will ask you to name an option. You could then proceed as follows:

```
OPTION? 10
TYPE IN THE SIZE OF THE MATRIX: N = ? 3
OPTION? 11
TYPE IN THE MATRIX A, ONE ROW AT A TIME, SEPARATING THE
ELEMENTS BY COMMAS.
ROW 1? 6.7, 3, 2.1
ROW 2? 4.5, -5.1, 0.5
ROW 3? 2.3, -5.6, 9.5
OPTION? 20
DIRECT ITERATION
OPTION? 30
TYPE IN THE ELEMENTS OF Y0, SEPARATING THE ELEMENTS BY
COMMAS:
ELEMENTS? 1, 1, 1
OPTION? 33
SIGNIFICANT FIGURES? 6
OPTION? 40
OUTLINE PRINTOUT
AFTER HOW MANY ITERATIONS DO YOU REQUIRE EACH
PRINTOUT? 5
OPTION? SOLVE
```

The computer would then respond with the printout given in the television section (printout (i) on page 25).

4.2 A list of options in EIGSOL

The options in this package are described under the same general headings as those in previous packages.

Command options

As before, these options have names, rather than numbers associated with them.

OPTIONS — this option prints the list of options available to you in this package.

SOLVE — this option computes the solution. It is the last option you use when solving a problem. Before this is used, all the data and the method of solution must have been given to the computer. If any of this information is missing, the computer will print an error message.

HELP — help may be obtained if you need it, at any stage that a question mark appears. If you type HELP after the computer has printed OPTION?, you will be asked questions to help you to decide what option to use next. These questions require an answer YES (Y) or NO (N).

If you type HELP while using an option, you will be given help as to how to use the option properly.

LIST — this option prints out the current information held in the computer about the current problem. It is advisable to ask for LIST before you ask for the problem to be solved. In this way, you can check that the computer is holding the correct information.

STOP — this option is your means of exit from the package. You should only use this option if

(i) you want to use another package, or

(ii) you want to log-off the computer.

The rest of the options in this package are obtained by typing a number when the computer types OPTION?

Problem options

10 Specify the size of the matrix

You will use option 10 to type in the size of the matrix. As all the matrices are square in this unit, it is only necessary to type in one number. For example, suppose you want to type in a 2×2 matrix. When the message

TYPE IN THE SIZE OF THE MATRIX: N = ?

appears, just type 2. The maximum size matrix you can use in this package is 10×10.

11 Enter matrix A

You will use option 11 to put the matrix **A** into the computer after you have specified the size of the matrix by using option 10. This is done in exactly the same way as in the computer package SIMLIN in *Unit 9*. When the message

ROW 1 ?

appears, type in the first row of the matrix, separating the elements by commas. If you type too few elements, an error message and an instruction to retype the whole row will be printed. If you type too many elements, a warning only will be printed.

12 Edit A

As in SIMLIN, this option is used to change individual elements of **A**. When you use this option, only the elements you want to alter will be changed.

To remind you how to use this option, here is a computer dialogue which will change the element a_{23} to be 8.

OPTION? 12

CHANGE ELEMENT IN ROW? 2

AND COLUMN? 3

TO HAVE THE VALUE? 8

DO YOU WANT TO CHANGE ANOTHER ELEMENT IN A? NO

OPTION?

Had you answered YES to the last question, you would have been asked to specify the next element you wanted to change.

13 **Choose a standard problem**

To save you having to type in some of the larger problems at the end of this section and in your tutor-marked assignment, there is a set of problems already stored in the computer. As in SIMLIN, each stored problem has a name. The following computer dialogue will access Exercise 5 (called PROBI) at the end of this section.

OPTION? 13

PROBLEM NAME? PROB 1

OPTION? LIST

It is advisable to print out the standard problem (using LIST) to check that the data is what you expected.

Method options

20 **Direct iteration**

If you use this option, the eigenvalue of largest modulus of **A** and the corresponding eigenvector of **A**, will be found. To do this, direct iteration as described in Procedure 2.4 is used.

See the notes on the use of these options in the next subsection.

To implement direct iteration you should specify the following parameters:

(i) the initial vector $\mathbf{y}_0$ (using Option 30);

(ii) the accuracy which you require in your answer (using Option 33).

21 **Inverse iteration**

If you use this option, the eigenvalue nearest to a specified value p, and the corresponding eigenvector of **A**, will be found. To do this, inverse iteration as described in Procedure 2.5 is used. When using this option, you should specify the following parameters:

(i) the initial vector $\mathbf{y}_0$ (using Option 30);

(ii) the accuracy which you require in your answer (using Option 33);

(iii) the value of p (using Option 32).

22 **LR method**

This option should be used when all the eigenvalues of a matrix $\mathbf{A} + q\mathbf{I}$ are required. The method used is the LR method described in Procedure 3.3. To use this option, the following values are specified:

(i) the value of q (using Option 34);

(ii) a condition for stopping the iteration (using Option 35).

23 **'Black Box' method**

This is an option provided for your convenience. Used in conjunction with Options 10 and 11, or with a standard problem (Option 13), this option will provide you with the eigenvalues and eigenvectors of any given matrix **A**, (both

real and complex). It is called a 'black box', simply because you do not know what is in it. The method used is a decomposition method, although not the one described in this unit.

You might like to use this option:

(i) to check the answers you get using the other methods (it replaces the option ANSWER in the other packages), or

(ii) to find quickly any eigenvalues and eigenvectors needed for the problems in *Units 22* and *24*.

Parameter options

30 **Enter $\mathbf{y}_0$**

You will use this option to specify a starting value for $\mathbf{y}_0$, both for the direct iteration method (Option 20) and the inverse iteration method (Option 21). When the message

ELEMENTS ?

is printed, type the elements of $\mathbf{y}_0$, separating the elements by commas. Typing too few elements will produce an error message, and an instruction to retype the whole vector. Typing too many elements will just be followed by a warning message.

If you fail to specify $\mathbf{y}_0$, then the first time either Option 20 or 21 is used, $\mathbf{y}_0$ will automatically be set to $[1 \quad 1 \quad \ldots \quad 1]^T$. After this, it will retain the value it had last time you used it. For instance, if you had set $\mathbf{y}_0$ to be $[1 \quad 1 \quad 0]^T$ in the previous problem, it will continue to hold this value in the present problem unless you change it.

31 **Edit $\mathbf{y}_0$**

You will use this option if you want to change individual elements in $\mathbf{y}_0$. Suppose you wished to change the second element of $\mathbf{y}_0$ to be 8.4. The computer dialogue to do this would go as follows:

OPTION? 31

CHANGE ELEMENT IN ROW? 2

TO HAVE THE VALUE? 8.4

DO YOU WANT TO CHANGE ANOTHER ELEMENT? NO

OPTION?

32 **Enter p**

This option is used to specify the value of p in $(\mathbf{A} - p\mathbf{I})^{-1}$ in the inverse iteration method (Option 21). When the message

P = ?

is printed, type in the value you choose for p. If you fail to specify p, then the first time Option 21 is used, p will be set to zero. After this, p will retain the last value you gave it—so be careful.

33 **Specify the accuracy required**

When you use this option in conjunction with Option 20 or 21, it specifies the degree of accuracy that you would like $\mathbf{y}_r$ to have. Suppose you would like two successive approximations to agree to five significant figures. Then, when the message

SIGNIFICANT FIGURES ?

is printed, you type 5. In this case, the iterations will continue until the corresponding elements in $\mathbf{y}_{r+1}$ and $\mathbf{y}_r$ agree to five significant figures.

If you fail to specify this option initially, the iteration will continue until the corresponding elements of $\mathbf{y}_{r+1}$ and $\mathbf{y}_r$ agree to three significant figures. After this, the accuracy will retain the last value you gave it. You may specify a maximum of eight significant figures.

34 Enter q

This option is used to specify the value of q when finding the eigenvalues of $\mathbf{A} + q\mathbf{I}$ using the LR method. Again, if q has never been specified, initially it will have the value zero. After this, q will retain the last value you gave it.

35 Stopping criterion for the LR method

When you use this option in conjunction with Option 22, it will stop the iterations when $\mathbf{A}_r$ has got sufficiently near upper triangular form. When the message

SUBDIAGONAL ELEMENTS LESS THAN ?

appears, if you type 0.0001, then the iterations will stop when all the elements underneath the main diagonal are less than 0.0001 in magnitude.

This does not in fact give you much indication as to how accurate your answer will be, but obviously, the smaller the value you give this parameter, the greater the accuracy you will achieve in your answer.

If you fail to specify this option initially, then the iteration will continue until all the elements underneath the main diagonal are less than 0.0001 in magnitude. The smallest value you are allowed to specify in this option is 10^{-6}.

Print options

40 Outline printout

This option gives a printout after each rth iteration. When the message

AFTER HOW MANY ITERATIONS DO YOU REQUIRE EACH PRINTOUT ?

appears, type the value of r. If you type 5, for example, every fifth iteration in the method of your choice will be printed.

41 Print solution only

This option will give the solution, and the number of iterations needed to satisfy the accuracy restrictions specified; no intermediate results will be printed. If no print option is specified this option will automatically be used.

4.3 Notes on the use of these options

Use of LIST

As you have a large choice of method options in this package, *you are strongly recommended to request a listing of the problem you are about to solve*, before you ask for the problem to be solved. This is to see that values for y_0, p, the number of significant figures etc., are all stored correctly before you attempt to solve the problem.

Use of Options 20, 21, 22

As the number of iterations may be quite large to achieve a sufficiently accurate answer, it is advisable to run the option once with 'solution only' (Option 41) to start with. This will tell you how many iterations were needed to achieve this answer, and will help you decide how much detail you would like to see using the more detailed printout (Option 40). This advice is given, as printout is very time consuming.

Use of Option 21

If inverse iteration is used with $p = 0$, then the routine simply becomes that for finding the eigenvalue of smallest modulus of $\mathbf{A}$.

Use of Option 22

(i) If all the eigenvalues of $\mathbf{A}$ are required, start with $q = 0$. Only in the event of the LR method not working will you want to try different values of q.

(ii) It is particularly important to use 'solution only' (Option 41) to begin with, for then a suitable message will be printed together with advice to alter the value of q if the method is not working.

(iii) If you require the eigenvectors as well as all the eigenvalues, find all the eigenvalues using the LR method and then use each eigenvalue in turn as your choice for p in the inverse iteration method (Option 21).

Use of option 23

This option can only be used with the 'solution only' option (Option 41), as the intermediate steps are not available for printing.

4.4 Computer exercises

You are expected to use the computer package EIGSOL to help you to solve the following problems. Some comments on the solutions to Exercises 1 and 3 are given on p. 50.

Exercise 1

(i) Use direct iteration to obtain the eigenvalue of largest modulus and the corresponding eigenvector for the matrix used on television:

$$\begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}.$$

Use the initial vectors

(a) $\mathbf{y}_0 = \begin{bmatrix} 3 \\ 1 \end{bmatrix}$ (b) $\mathbf{y}_0 = \begin{bmatrix} -2 \\ 1 \end{bmatrix}$.

Can you explain what has happened?

(ii) Now use direct iteration again, with the initial vectors

(a) $\mathbf{y}_0 = \begin{bmatrix} -2 \\ 1.01 \end{bmatrix}$, (b) $\mathbf{y}_0 = \begin{bmatrix} -2 \\ 1.00001 \end{bmatrix}$, (c) $\mathbf{y}_0 = \begin{bmatrix} -2 \\ 0.99999 \end{bmatrix}$,

(d) $\mathbf{y}_0 = \begin{bmatrix} -2.00001 \\ 1 \end{bmatrix}$.

Explain the results you now obtain in light of the original two results.

Exercise 2

Use the computer package to find all the eigenvalues, and corresponding eigenvectors, of

(i) $\begin{bmatrix} 2 & 3 & 2 \\ 10 & 3 & 4 \\ 3 & 6 & 1 \end{bmatrix}$, (ii) $\begin{bmatrix} 6.1 & -2.3 & 3.2 \\ 1 & -3.4 & 9.7 \\ 3.5 & 7.1 & 11 \end{bmatrix}$.

In all cases start with an initial vector $\mathbf{y}_0 = [1 \quad 1 \quad 1]^T$, and continue until two consecutive iterates agree to six significant figures. Proceed as follows:

(a) Use direct iteration to obtain the eigenvalue of largest modulus.

(b) Use inverse iteration to find the eigenvalue of smallest modulus.

(c) Suggest a suitable value of p, for use in the inverse iteration method to find the remaining eigenvalue, and hence find the remaining eigenvalue.

Check your results using Option 23.

Exercise 3

(i) Use the LR method to find the eigenvalues of the matrices in Exercise 2. Specify six significant figure accuracy using Option 35.

Your answers may differ slightly from those in Exercise 2. Can you explain this?

(ii) Use the eigenvalues you have just found as suitable values of p in the inverse iteration method to obtain the corresponding eigenvectors.

Exercise 4

(i) Check that the LR method fails for the matrix

$$\mathbf{A} = \begin{bmatrix} 0 & 2 & 3 \\ 2 & 3 & 0 \\ 1 & 7 & 10 \end{bmatrix}.$$

(ii) Use the LR method to find the eigenvalues of a matrix $\mathbf{A} + q\mathbf{I}$, with a suitable choice of q. Hence find the eigenvalues of $\mathbf{A}$.

Exercise 5

(To save you time, the coefficients of this problem are already stored in the computer, under the name PROBI. *Use Option* 13 *to obtain* PROBI.*)*

Find the eigenvalue

(i) of largest modulus

(ii) of least modulus

of the matrix

$$\mathbf{A} = \begin{bmatrix} 12.4 & 1.7 & 3.5 & 1.1 & -2.7 \\ 1.7 & 10.8 & -1.1 & 2.7 & 3.1 \\ 3.5 & -1.1 & 7.8 & -3.0 & 0.1 \\ 1.1 & 2.7 & -3.0 & 8.4 & 1.4 \\ -2.7 & 3.1 & 0.1 & 1.4 & 9.6 \end{bmatrix}.$$

(iii) Use trial values of p in the inverse iteration method to find the other eigenvalues of $\mathbf{A}$.

(iv) Use the LR method to check the eigenvalues which you have found.

Summary of Section 4

This section describes the options in the computer package EIGSOL. This package implements the methods described in Sections 2 and 3, all of which find eigenvalues of a matrix $\mathbf{A}$.

The following lists of options are the minimum needed to implement these methods.

Description of option	Direct iteration	Inverse iteration	LR method
Specify the size of $\mathbf{A}$	10	10	10
Enter $\mathbf{A}$	11	11	11
Method to be used	20	21	22
Specify p	not applicable	32	not applicable
Specify q	not applicable	not applicable	34
Print solution only	41	41	41
Advisable to list before solving	LIST	LIST	LIST
Solve the problem	SOLVE	SOLVE	SOLVE

5 End of unit test

The first four exercises cover routine material from the unit and you should make sure you can tackle them with confidence. The remaining exercises may require a little more thought.

The first three exercises all refer to the matrix

$$\mathbf{A} = \begin{bmatrix} 7 & 14 \\ 3 & 8 \end{bmatrix}.$$

Exercise 1

Find, directly, the eigenvalues and eigenvectors of $\mathbf{A}$.

Exercise 2

Use suitable iterative methods to find approximations to the following selected eigenvalues and corresponding eigenvectors of $\mathbf{A}$:

(i) the eigenvalue of largest modulus,

(ii) the eigenvalue of smallest modulus,

(iii) the eigenvalue closest to 12.

In all cases, use a starting vector $\mathbf{y}_0 = [1 \quad 1]^T$, and continue the iterations until the elements in two iterates $\mathbf{y}_{r+1}$ and $\mathbf{y}_r$ agree to two significant figures.

There are, of course, only two eigenvalues and so the iterations in (iii) will converge to one of the eigenvalues found in (i) and (ii). The purpose of (iii) is to give you extra practice.

Exercise 3

Use three iterations of the LR method to make an approximation of the eigenvalues of $\mathbf{A}$.

Exercise 4

Given that $\mathbf{A}$ is an $n \times n$ matrix with distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$, and corresponding eigenvectors $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$, find the eigenvalues and eigenvectors of the following matrices in terms of the eigenvalues and eigenvectors of $\mathbf{A}$:

(i) $\mathbf{A}^3$

(ii) $(\mathbf{A} + p\mathbf{I})^{-1}(\mathbf{A} - p\mathbf{I})$

Exercise 5

In certain cases, symmetric matrices have a decomposition $\mathbf{LL}^T$, where $\mathbf{L}$ is a lower triangular matrix. For instance, the 2×2 matrix $\begin{bmatrix} a & b \\ b & c \end{bmatrix}$ has the decomposition

$$\begin{bmatrix} \sqrt{a} & 0 \\ \dfrac{b}{\sqrt{a}} & \sqrt{\dfrac{\det \mathbf{A}}{a}} \end{bmatrix} \begin{bmatrix} \sqrt{a} & \dfrac{b}{\sqrt{a}} \\ 0 & \sqrt{\dfrac{\det \mathbf{A}}{a}} \end{bmatrix}.$$

so long as $a > 0$ and $\det \mathbf{A} > 0$. This type of decomposition is called the **Cholesky decomposition**.

(i) Show that $\mathbf{L}^T\mathbf{L}$ is also a symmetric matrix.

(ii) Find the Cholesky decomposition of the matrix

$$\mathbf{A} = \begin{bmatrix} 9 & 8 \\ 8 & 9 \end{bmatrix}.$$

Exercise 6

Consider the following iterative scheme:

Given a symmetric matrix $\mathbf{A}_0$

1. Find the Cholesky decomposition

$$\mathbf{A}_r = \mathbf{L}_r\mathbf{L}_r^T.$$

2. Form the new symmetric matrix

$$\mathbf{A}_{r+1} = \mathbf{L}_r^T\mathbf{L}_r.$$

Do 1 and 2 for $r = 0, 1, 2 \ldots$

In this way a series of symmetric matrices $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2, \ldots$ are formed.

(i) Show that the sequence of matrices formed all have the same eigenvalues.

(ii) Find $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{A}_3$ for

$$\mathbf{A}_0 = \begin{bmatrix} 9 & 8 \\ 8 & 9 \end{bmatrix}.$$

Hence, find an approximation for the eigenvalues of $\mathbf{A}_0$.

Exercise 7

Suppose that a given matrix $\mathbf{A}$ has both an LU decomposition and a Cholesky decomposition, and that both the LR method and the method described in Exercise 6 work. Which method would you use to find an approximation to the eigenvalues of $\mathbf{A}$, and why?

[*Solutions to Exercises 1 to 7 on pp. 50–52*]

Appendix: Solutions to the exercises

Solutions to the exercises in Section 1

1. **(i)** For the matrix

$$\mathbf{A} = \begin{bmatrix} 7 & 3 \\ 3 & 7 \end{bmatrix}$$

the condition $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ gives

$$\begin{vmatrix} 7-\lambda & 3 \\ 3 & 7-\lambda \end{vmatrix} = 0.$$

Hence we obtain the equation

$$(7-\lambda)^2 - 3^2 = 0$$

or $\quad (4-\lambda)(10-\lambda) = 0$

and so the eigenvalues of $\mathbf{A}$ are 4 and 10.

(ii) For the matrix

$$\mathbf{A} = \begin{bmatrix} 3 & 2 \\ 1 & 4 \end{bmatrix}$$

$\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ gives

$$\begin{vmatrix} 3-\lambda & 2 \\ 1 & 4-\lambda \end{vmatrix} = 0.$$

Hence we obtain the equation

$$(3-\lambda)(4-\lambda) - 2 = 0.$$

This gives us

$$\lambda^2 - 7\lambda + 10 = 0$$

or $\quad (\lambda-5)(\lambda-2) = 0.$

Hence, the eigenvalues of $\mathbf{A}$ are 5 and 2.

2. **(i)** The equations determining the eigenvectors are

$$(7-\lambda)x_1 + 3x_2 = 0.$$
$$3x_1 + (7-\lambda)x_2 = 0.$$

We shall consider each eigenvalue in turn.

Case $\lambda = 4$: in this case the equations become

$$3x_1 + 3x_2 = 0. \qquad \text{(twice)}$$

Thus we get a solution $x_1 = -k$, $x_2 = k$. So, an eigenvector corresponding to $\lambda = 4$ is $\begin{bmatrix} -1 \\ 1 \end{bmatrix}$.

Case $\lambda = 10$: in this case the equations become

$$-3x_1 + 3x_2 = 0$$
$$3x_1 - 3x_2 = 0.$$

(Note that these equations provide us with the same information.)

These have a solution of the form $x_1 = k$, $x_2 = k$. So, an eigenvector corresponding to $\lambda = 10$ is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

(ii) The equations determining the eigenvectors are

$$(3-\lambda)x_1 + 2x_2 = 0$$
$$x_1 + (4-\lambda)x_2 = 0.$$

We shall consider each eigenvalue in turn.

Case $\lambda = 2$: in this case the equations become

$$x_1 + 2x_2 = 0$$
$$x_1 + 2x_2 = 0.$$

Putting $x_1 = k$, we get $x_2 = -\frac{1}{2}k$. So an eigenvector corresponding to $\lambda = 2$ is $\begin{bmatrix} 1 \\ -\frac{1}{2} \end{bmatrix}$.

Case $\lambda = 5$: in this case the equations become

$$-2x_1 + 2x_2 = 0$$
$$x_1 - x_2 = 0.$$

These equations have a solution of the form $x_2 = k$, $x_1 = k$. So an eigenvector corresponding to $\lambda = 5$ is $\begin{bmatrix} 1 \\ 1 \end{bmatrix}$.

3. The condition $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ gives us

$$\begin{vmatrix} 1-\lambda & 2 & 6 \\ 0 & 4-\lambda & 4 \\ 0 & 0 & 2-\lambda \end{vmatrix} = 0.$$

Hence $(1-\lambda)(4-\lambda)(2-\lambda) = 0$, and the eigenvalues are 1, 4, and 2. To find the corresponding eigenvectors, we solve the equation

$$\begin{bmatrix} 1-\lambda & 2 & 6 \\ 0 & 4-\lambda & 4 \\ 0 & 0 & 2-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0},$$

putting λ equal to each eigenvalue in turn.

Case $\lambda = 1$: in this case we obtain the equations

$$2x_2 + 6x_3 = 0$$
$$3x_2 + 4x_3 = 0$$
$$x_3 = 0.$$

This gives us $x_3 = 0$, $x_2 = 0$. As x_1 is not involved in these equations, it can take any value k.

Hence $[1 \quad 0 \quad 0]^T$ is an eigenvector corresponding to $\lambda = 1$.

Case $\lambda = 4$: in this case we obtain the equations

$$-3x_1 + 2x_2 + 6x_3 = 0$$
$$x_3 = 0 \qquad \text{(twice).}$$

Thus

$$-3x_1 + 2x_2 = 0,$$

and putting $x_1 = k$, we get $x_2 = \frac{3}{2}k$.

So $[1 \quad \frac{3}{2} \quad 0]^T$ (or $[2 \quad 3 \quad 0]^T$) is an eigenvector corresponding to $\lambda = 4$.

Case $\lambda = 2$: in this case we obtain the equations

$$-x_1 + 2x_2 + 6x_3 = 0$$
$$2x_2 + 4x_3 = 0.$$

So putting $x_3 = k$ we find that $x_2 = -2k$ and $x_1 = 2k$. So $[2 \quad -2 \quad 1]^T$ is an eigenvector corresponding to $\lambda = 2$.

4. From Example 7 we know that the eigenvalues of $\mathbf{B}$ are $1 + 2i$ and $1 - 2i$. To find the corresponding eigenvectors, we solve $(\mathbf{B} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, that is

$$\begin{bmatrix} 1-\lambda & -1 \\ 4 & 1-\lambda \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0},$$

putting λ equal to each eigenvalue in turn.

Case $\lambda = 1 + 2i$: in this case we obtain the equations

$$-2ix_1 - x_2 = 0$$
$$4x_1 - 2ix_2 = 0.$$

(Note that the second equation is simply $2i$ times the first one.) From either equation, we obtain solutions $x_1 = ik/2$, $x_2 = k$.

Hence, an eigenvector corresponding to $1 + 2i$ is $\begin{bmatrix} i/2 \\ 1 \end{bmatrix}$.

Case $\lambda = 1 - 2i$: in this case we obtain the equations

$$\begin{aligned} 2ix_1 - x_2 &= 0 \\ 4x_1 + 2ix_2 &= 0. \end{aligned}$$

From either equation, we obtain solutions of the form $x_1 = -ik/2$, $x_2 = k$.

Hence, an eigenvector corresponding to $1 - 2i$ is $\begin{bmatrix} -i/2 \\ 1 \end{bmatrix}$.

5. (i) The characteristic equation is given by

$$\begin{vmatrix} 1-\lambda & 2 \\ -1 & 4-\lambda \end{vmatrix} = 0,$$

that is

$$(1-\lambda)(4-\lambda) + 2 = 0$$

or $\lambda^2 - 5\lambda + 6 = 0$

i.e. $(\lambda - 2)(\lambda - 3) = 0.$

So the eigenvalues are 2 and 3. To obtain the eigenvectors, we solve the equation

$$\begin{bmatrix} 1-\lambda & 2 \\ -1 & 4-\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

putting λ equal to each eigenvalue in turn.

Case $\lambda = 2$: we obtain the equations

$$\begin{aligned} -x_1 + 2x_2 &= 0 \\ -x_1 + 2x_2 &= 0 \end{aligned}$$

which have solutions of the form $x_1 = 2k$, $x_2 = k$. Hence $[2 \quad 1]^T$ is an eigenvector corresponding to the eigenvalue 2.

Case $\lambda = 3$: we obtain the equations

$$\begin{aligned} -2x_1 + 2x_2 &= 0 \\ -x_1 + x_2 &= 0 \end{aligned}$$

which have solutions of the form $x_1 = k$, $x_2 = k$. Hence $[1 \quad 1]^T$ is an eigenvector corresponding to the eigenvalue 3.

(ii) The characteristic equation is given by

$$\begin{vmatrix} 4-\lambda & 2 \\ 5 & 7-\lambda \end{vmatrix} = 0,$$

that is

$$(4-\lambda)(7-\lambda) - 10 = 0$$

or $\lambda^2 - 11\lambda + 18 = 0$

i.e. $(\lambda - 9)(\lambda - 2) = 0.$

So, the eigenvalues are 9 and 2. To obtain the eigenvectors, we solve the equation

$$\begin{bmatrix} 4-\lambda & 2 \\ 5 & 7-\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

putting λ equal to each eigenvalue in turn.

Case $\lambda = 9$: we obtain the equations

$$\begin{aligned} -5x_1 + 2x_2 &= 0 \\ 5x_1 - 2x_2 &= 0 \end{aligned}$$

which have solutions of the form $x_2 = k$, $x_1 = \frac{2}{5}k$. Hence, $[2 \quad 5]^T$ is an eigenvector corresponding to the eigenvalue 9.

Case $\lambda = 2$: we obtain the equations

$$\begin{aligned} 2x_1 + 2x_2 &= 0 \\ 5x_1 + 5x_2 &= 0 \end{aligned}$$

which have solutions of the form $x_1 = k$, $x_2 = -k$. Hence $[1 \quad -1]^T$ is an eigenvector corresponding to the eigenvalue 2.

(iii) The characteristic equation is given by

$$\begin{vmatrix} 1-\lambda & -1 \\ 1 & 1-\lambda \end{vmatrix} = 0,$$

that is,

$$(1-\lambda)^2 + 1 = 0,$$

or $(1 - \lambda + i)(1 - \lambda - i) = 0.$

so, the eigenvalues are $1 + i$ and $1 - i$. To obtain the eigenvectors, we solve the equations

$$\begin{bmatrix} 1-\lambda & -1 \\ 1 & 1-\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \end{bmatrix} = \mathbf{0}$$

putting λ equal to each eigenvalue in turn.

Case $\lambda = 1 + i$: we obtain the equations

$$\begin{aligned} -ix_1 - x_2 &= 0 \\ x_1 - ix_2 &= 0, \end{aligned}$$

which have solutions of the form $x_1 = ik$, $x_2 = k$. Hence, $[i \quad 1]^T$ is an eigenvector corresponding to the eigenvalue $1 + i$.

Case $\lambda = 1 - i$: we obtain the equations

$$\begin{aligned} ix_1 - x_2 &= 0 \\ x_1 + ix_2 &= 0, \end{aligned}$$

which have solutions of the form $x_1 = -ik$, $x_2 = k$. Hence $[-i \quad 1]^T$ is an eigenvector corresponding to the eigenvalue $1 - i$.

(iv) The characteristic equation is given by

$$\begin{vmatrix} 1-\lambda & 0 & -1 \\ 1 & 2-\lambda & 1 \\ 2 & 2 & 3-\lambda \end{vmatrix} = 0.$$

Expanding the determinant, we obtain the equation

$$(1-\lambda)\begin{vmatrix} 2-\lambda & 1 \\ 2 & 3-\lambda \end{vmatrix} - \begin{vmatrix} 1 & 2-\lambda \\ 2 & 2 \end{vmatrix} = 0.$$

This gives

$$(1-\lambda)(2-\lambda)(3-\lambda) - 2(1-\lambda) - (2 - 2(2-\lambda)) = 0$$

which simplifies to give

$$(1-\lambda)(2-\lambda)(3-\lambda) = 0.$$

So the eigenvalues are 1, 2 and 3. To obtain the eigenvectors, we solve the equation

$$\begin{bmatrix} 1-\lambda & 0 & -1 \\ 1 & 2-\lambda & 1 \\ 2 & 2 & 3-\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

putting λ equal to each eigenvalue in turn.

Case $\lambda = 1$: we obtain the equations

$$\begin{aligned} -x_3 &= 0 \\ x_1 + x_2 + x_3 &= 0 \\ 2x_1 + 2x_2 + 2x_3 &= 0 \end{aligned}$$

which have solutions of the form $x_1 = -k$, $x_2 = k$, $x_3 = 0$. Hence $[-1 \quad 1 \quad 0]^T$ is an eigenvector corresponding to the eigenvalue 1.

Case $\lambda = 2$: we obtain the equations

$$\begin{aligned} -x_1 \phantom{{}+2x_2} - x_3 &= 0 \\ x_1 \phantom{{}+2x_2} + x_3 &= 0 \\ 2x_1 + 2x_2 + x_3 &= 0 \end{aligned}$$

which have solutions of the form $x_3 = k$, $x_1 = -k$, and $x_2 = k/2$.

Hence, $[-2 \quad 1 \quad 2]^T$ is an eigenvector corresponding to the eigenvalue 2.

Case $\lambda = 3$: we obtain the equations

$$\begin{aligned} -2x_1 \qquad - x_3 &= 0 \\ x_1 - x_2 + x_3 &= 0 \\ 2x_1 + 2x_2 \qquad &= 0 \end{aligned}$$

which have solutions of the form $x_3 = k$, $x_1 = -k/2$, $x_2 = k/2$.

Hence $[-1 \quad 1 \quad 2]^T$ is an eigenvector corresponding to the eigenvalue 3.

(v) The characteristic equation is given by

$$\begin{vmatrix} 3-\lambda & 2 & 2 \\ 2 & 2-\lambda & 0 \\ 2 & 0 & 4-\lambda \end{vmatrix} = 0.$$

Expanding the determinant, we obtain the equation

$$(3-\lambda)\begin{vmatrix} 2-\lambda & 0 \\ 0 & 4-\lambda \end{vmatrix} - 2\begin{vmatrix} 2 & 0 \\ 2 & 4-\lambda \end{vmatrix} + 2\begin{vmatrix} 2 & 2-\lambda \\ 2 & 0 \end{vmatrix} = 0.$$

Hence

$$(3-\lambda)(2-\lambda)(4-\lambda) - 4(4-\lambda) - 4(2-\lambda) = 0$$

so $$(3-\lambda)(2-\lambda)(4-\lambda) - 8(3-\lambda) = 0$$

i.e. $$(3-\lambda)(\lambda^2 - 6\lambda) = 0.$$

Thus, we obtain the eigenvalues 0, 3, and 6. To find the corresponding eigenvectors we solve

$$\begin{bmatrix} 3-\lambda & 2 & 2 \\ 2 & 2-\lambda & 0 \\ 2 & 0 & 4-\lambda \end{bmatrix}\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \mathbf{0}$$

putting λ equal to each eigenvalue in turn.

Case $\lambda = 0$: we obtain the equations

$$\begin{aligned} 3x_1 + 2x_2 + 2x_3 &= 0 \\ 2x_1 + 2x_2 \qquad &= 0 \\ 2x_1 \qquad + 4x_3 &= 0, \end{aligned}$$

which have solutions of the form $x_3 = k$, $x_1 = -2k$, $x_2 = 2k$.

Hence, $[-2 \quad 2 \quad 1]^T$ is an eigenvector corresponding to the eigenvalue 0.

Case $\lambda = 3$: we obtain the equations

$$\begin{aligned} 2x_2 + 2x_3 &= 0 \\ 2x_1 - x_2 \qquad &= 0 \\ 2x_1 \qquad + x_3 &= 0 \end{aligned}$$

which have solutions of the form $x_3 = k$, $x_2 = -k$, $x_1 = -k/2$.

Hence, $[1 \quad 2 \quad -2]^T$ is an eigenvector corresponding to the eigenvalue 3.

Case $\lambda = 6$: we obtain the equations

$$\begin{aligned} -3x_1 + 2x_2 + 2x_3 &= 0 \\ 2x_1 - 4x_2 \qquad &= 0 \\ 2x_1 \qquad - 2x_3 &= 0 \end{aligned}$$

which have solutions of the form $x_3 = k$, $x_1 = k$, $x_2 = k/2$.

Hence, $[2 \quad 1 \quad 2]^T$ is an eigenvector corresponding to the eigenvalue 6.

Solutions to the exercises in Section 2

1. We have seen that for any number q, the matrix $(\mathbf{A} + q\mathbf{I})$ has eigenvalues $\lambda_1 + q, \ \lambda_2 + q, \ \ldots, \ \lambda_n = q$, so

(i) $\mathbf{A} + 5\mathbf{I}$ has eigenvalues $\lambda_1 + 5, \ \lambda_2 + 5, \ \ldots, \ \lambda_n + 5$.

(ii) $\mathbf{A} - 2\mathbf{I}$ has eigenvalues $\lambda_1 - 2, \ \lambda_2 - 2, \ \ldots, \ \lambda_n - 2$.

(iii) $\mathbf{A} - q\mathbf{I}$ has eigenvalues $\lambda_1 - q, \ \lambda_2 - q, \ \ldots, \ \lambda_n - q$.

2. We know

$$\mathbf{A}\mathbf{x}_i = \lambda_i\mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n.$$

Subtracting $p\mathbf{x}_i$ (or equivalently $p\mathbf{I}\mathbf{x}_i$) from both sides of this equation, we get

$$\mathbf{A}\mathbf{x}_i - p\mathbf{I}\mathbf{x}_i = \lambda_i\mathbf{x}_i - p\mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n,$$

so $$(\mathbf{A} - p\mathbf{I})\mathbf{x}_i = (\lambda_i - p)\mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n.$$

Since $(\mathbf{A} - p\mathbf{I})$ is non-singular, we can multiply both sides of this equation by $(\mathbf{A} - p\mathbf{I})^{-1}$ to obtain

$$\mathbf{x}_i = (\lambda_i - p)(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n,$$

or $$(\mathbf{A} - p\mathbf{I})^{-1}\mathbf{x}_i = \frac{1}{\lambda_i - p}\mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n.$$

Thus $(\mathbf{A} - p\mathbf{I})^{-1}$ has the set of eigenvalues $\left\{\frac{1}{\lambda_1 - p}, \frac{1}{\lambda_2 - p}, \ldots, \frac{1}{\lambda_n - p}\right\}$, and has the same eigenvectors as $\mathbf{A}$.

3. *Step 1*:

$$\mathbf{y}_1' = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 12 \end{bmatrix}.$$

Dividing by 12 gives $\mathbf{y}_1 = [0.5 \quad 1]^T$.

Step 2:

$$\mathbf{y}_2' = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix}\begin{bmatrix} 0.5 \\ 1 \end{bmatrix} = \begin{bmatrix} 4 \\ 9.5 \end{bmatrix}.$$

Dividing by 9.5 gives $\mathbf{y}_2 = [0.421 \quad 1]^T$.

Step 3:

$$\mathbf{y}_3' = \begin{bmatrix} 4 & 2 \\ 5 & 7 \end{bmatrix}\begin{bmatrix} 0.421 \\ 1 \end{bmatrix} = \begin{bmatrix} 3.684 \\ 9.105 \end{bmatrix}.$$

Dividing by 9.105 gives $\mathbf{y}_3 = [0.405 \quad 1]^T$.

Thus we obtain an estimate of 9.105 for the eigenvalue of largest modulus and a corresponding eigenvector $[0.405 \quad 1]^T$.

These compare well with the actual value 9 and $[0.4 \quad 1]^T$.

4. Given $\mathbf{A} = \begin{bmatrix} 7 & 3 \\ 8 & 5 \end{bmatrix}$, then $\mathbf{A}^{-1} = \frac{1}{11}\begin{bmatrix} 5 & -3 \\ -8 & 7 \end{bmatrix}$.

Using the scheme $\mathbf{y}_{r+1} = \mathbf{A}^{-1}\mathbf{y}_r$ with scaling, and an initial vector $\mathbf{y}_0 = [1 \quad 1]^T$ we obtain

Step 1:

$$\mathbf{y}_1' = \frac{1}{11}\begin{bmatrix} 5 & -3 \\ -8 & 7 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.182 \\ -0.091 \end{bmatrix}.$$

Dividing by 0.182 we obtain $\mathbf{y}_1 = [1 \quad -0.5]^T$.

Step 2:

$$\mathbf{y}_2' = \frac{1}{11}\begin{bmatrix} 5 & -3 \\ -8 & 7 \end{bmatrix}\begin{bmatrix} 1 \\ -0.5 \end{bmatrix} = \begin{bmatrix} 0.591 \\ -1.045 \end{bmatrix}.$$

Dividing by -1.045 we obtain $\mathbf{y}_2 = [-0.566 \quad 1]^T$.

Step 3:

$$\mathbf{y}_3' = \frac{1}{11}\begin{bmatrix} 5 & -3 \\ -8 & 7 \end{bmatrix}\begin{bmatrix} -0.566 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.530 \\ 1.047 \end{bmatrix}.$$

Dividing by 1.047 we obtain $\mathbf{y}_3 = [-0.506 \quad 1]^T$.

Thus, the estimated eigenvalue of smallest modulus for $\mathbf{A}$ is $1/1.047$, i.e. 0.955, with a corresponding estimated eigenvector $[-0.506 \quad 1]^T$. Although this is not a bad approximation to the actual eigenvalue 1 and eigenvector $[-0.5 \quad 1]^T$, until the values of $\mathbf{y}_r$ settle down a bit more it would not be very sensible to stop at this point in practice.

5. Given $\mathbf{A} = \begin{bmatrix} 2 & 1 \\ 1 & 4 \end{bmatrix}$ and $p = 2$, then $\mathbf{A} - p\mathbf{I} = \begin{bmatrix} 0 & 1 \\ 1 & 2 \end{bmatrix}$ and so

$$(\mathbf{A} - p\mathbf{I})^{-1} = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}.$$

Using the scheme $\mathbf{y}_{r+1} = (\mathbf{A} - p\mathbf{I})^{-1}\mathbf{y}_r$ with scaling, and an initial vector $\mathbf{y}_0 = [1 \quad 0]^T$ we obtain

Step 1:

$$\mathbf{y}_1' = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -2 \\ 1 \end{bmatrix}.$$

Dividing by -2 we obtain $\mathbf{y}_1 = [1 \quad -0.5]^T$.

Step 2:

$$\mathbf{y}_2' = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -0.5 \end{bmatrix} = \begin{bmatrix} -2.5 \\ 1 \end{bmatrix}.$$

Dividing by -2.5 we obtain $\mathbf{y}_2 = [1 \quad -0.4]^T$.

Step 3:

$$\mathbf{y}_3' = \begin{bmatrix} -2 & 1 \\ 1 & 0 \end{bmatrix}\begin{bmatrix} 1 \\ -0.4 \end{bmatrix} = \begin{bmatrix} -2.4 \\ 1 \end{bmatrix}.$$

Dividing by -2.4 we obtain $\mathbf{y}_3 = [1 \quad -0.417]^T$.

Hence, an estimate for the eigenvalue of $\mathbf{A}$ nearest to 2 is $\dfrac{1}{-2.4} + 2$ or 1.583, and an approximation to the corresponding eigenvector is $[1 \quad -0.417]^T$. The actual eigenvalues are 4.412 and 1.588 (to three decimal places). Although our approximation was not too bad, it would have been advisable to continue, as the process had not settled down sufficiently to quote this result with confidence.

6. (i) The direct iteration scheme $\mathbf{y}_{r+1} = \mathbf{A}\mathbf{y}_r$, with scaling gives

Step 1:

$$\mathbf{y}_1' = \begin{bmatrix} -39 & 40 \\ -20 & 21 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} -39 \\ -20 \end{bmatrix}.$$

Dividing by -39 gives $\mathbf{y}_1 = [1 \quad 0.513]^T$.

Step 2:

$$\mathbf{y}_2' = \begin{bmatrix} -39 & 40 \\ -20 & 21 \end{bmatrix}\begin{bmatrix} 1 \\ 0.513 \end{bmatrix} = \begin{bmatrix} -18.487 \\ -9.231 \end{bmatrix}.$$

Dividing by -18.487 gives $\mathbf{y}_2 = [1 \quad 0.499]^T$.

Step 3:

$$\mathbf{y}_3' = \begin{bmatrix} -39 & 40 \\ -20 & 21 \end{bmatrix}\begin{bmatrix} 1 \\ 0.499 \end{bmatrix} = \begin{bmatrix} -19.028 \\ -9.515 \end{bmatrix}.$$

Dividing by -19.028 gives $\mathbf{y}_3 = [1 \quad 0.500]^T$.

Hence an estimate for the eigenvalue of largest modulus is -19.028 and an estimate for a corresponding eigenvector is $[1 \quad 0.5]^T$.

(ii) With $p = 2$, $\mathbf{A} - p\mathbf{I} = \begin{bmatrix} -41 & 40 \\ -20 & 19 \end{bmatrix}$ and so

$$(\mathbf{A} - p\mathbf{I})^{-1} = \frac{1}{21}\begin{bmatrix} 19 & -40 \\ 20 & -41 \end{bmatrix}.$$

Using the inverse iteration scheme $\mathbf{y}_{r+1} = (\mathbf{A} - p\mathbf{I})^{-1}\mathbf{y}_r$, with scaling gives

Step 1:

$$\mathbf{y}_1' = \frac{1}{21}\begin{bmatrix} 19 & -40 \\ 20 & -41 \end{bmatrix}\begin{bmatrix} 1 \\ 0 \end{bmatrix} = \begin{bmatrix} 0.905 \\ 0.952 \end{bmatrix}.$$

Dividing by 0.952 gives $\mathbf{y}_1 = [0.95 \quad 1]^T$.

Step 2:

$$\mathbf{y}_2' = \frac{1}{21}\begin{bmatrix} 19 & -40 \\ 20 & -41 \end{bmatrix}\begin{bmatrix} 0.95 \\ 1 \end{bmatrix} = \begin{bmatrix} -1.045 \\ -1.048 \end{bmatrix}.$$

Dividing by -1.048 gives $\mathbf{y}_2 = [0.998 \quad 1]^T$.

Step 3:

$$\mathbf{y}_3' = \frac{1}{21}\begin{bmatrix} 19 & -40 \\ 20 & -41 \end{bmatrix}\begin{bmatrix} 0.998 \\ 1 \end{bmatrix} = \begin{bmatrix} -1.0021 \\ -1.0022 \end{bmatrix}.$$

Dividing by -1.0022 gives $\mathbf{y}_3 = [1.000 \quad 1]^T$.

Hence an estimate for the eigenvalue of $\mathbf{A}$ nearest to 2 is $\dfrac{1}{-1.0022} + 2 = 1.002$, and $[1 \quad 1]^T$ is an estimate for the corresponding eigenvector.

Solutions to the exercises in Section 3

1. We require

$$\begin{bmatrix} 6 & 5 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ a & 1 \end{bmatrix}\begin{bmatrix} b & c \\ 0 & d \end{bmatrix} = \begin{bmatrix} b & c \\ ab & ac + d \end{bmatrix}.$$

Hence

$$b = 6, \qquad c = 5,$$
$$ab = 4, \qquad ac + d = 5.$$

So $a = \frac{2}{3}$, and $d = \frac{5}{3}$.

Thus we obtain the decomposition

$$\begin{bmatrix} 6 & 5 \\ 4 & 5 \end{bmatrix} = \begin{bmatrix} 1 & 0 \\ \frac{2}{3} & 1 \end{bmatrix}\begin{bmatrix} 6 & 5 \\ 0 & \frac{5}{3} \end{bmatrix}.$$

2. We require

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 6 & 4 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ a & 1 & 0 \\ b & c & 1 \end{bmatrix}\begin{bmatrix} d & e & f \\ 0 & g & h \\ 0 & 0 & j \end{bmatrix}$$
$$= \begin{bmatrix} d & e & f \\ ad & ae + g & af + h \\ bd & be + cg & bf + ch + j \end{bmatrix}.$$

Equating coefficients, we obtain:

First row: $d = 1$, $e = 2$, $f = 3$.

First column (remainder!): $ad = 1$, $bd = 2$,

which gives $a = 1$ and $b = 2$.

Second row: $ae + g = 6$, $af + h = 4$,

which gives $g = 4$ and $h = 1$.

Second column: $be + cg = 1$,

which gives $c = -\frac{3}{4}$.

Third row: $bf + ch + j = 4$,

which gives $j = -\frac{5}{4}$.

Thus the decomposition is

$$\begin{bmatrix} 1 & 2 & 3 \\ 1 & 6 & 4 \\ 2 & 1 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 \\ 1 & 1 & 0 \\ 2 & -\frac{3}{4} & 1 \end{bmatrix}\begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 1 \\ 0 & 0 & -\frac{5}{4} \end{bmatrix}.$$

We can check this by multiplying out the right-hand side.

Note that there is more than one order in which you can equate the coefficients. The row and column approach used here is quite a standard one. You can equally well find the coefficients row by row, or column by column.

3. First consider **AB**

$$\mathbf{AB} = \begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix}\begin{bmatrix} 1 & 4 \\ 0 & -10 \end{bmatrix} = \begin{bmatrix} 1 & 4 \\ 3 & 2 \end{bmatrix}.$$

The eigenvalues of **AB** are given by $\det(\mathbf{AB} - \lambda\mathbf{I}) = 0$. That is

$$\begin{vmatrix} 1-\lambda & 4 \\ 3 & 2-\lambda \end{vmatrix} = 0$$

giving

$$(1-\lambda)(2-\lambda) - 12 = 0$$

or $\quad \lambda^2 - 3\lambda - 10 = 0$

i.e. $\quad (\lambda - 5)(\lambda + 2) = 0.$

Thus we obtain eigenvalues 5 and -2 for **AB**.

Now consider **BA**

$$\mathbf{BA} = \begin{bmatrix} 1 & 4 \\ 0 & -10 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 3 & 1 \end{bmatrix} = \begin{bmatrix} 13 & 4 \\ -30 & -10 \end{bmatrix}.$$

The eigenvalues of **BA** are given by $\det(\mathbf{BA} - \lambda\mathbf{I}) = 0$. That is

$$\begin{vmatrix} 13-\lambda & 4 \\ -30 & -10-\lambda \end{vmatrix} = 0$$

giving

$$(13-\lambda)(-10-\lambda) + 120 = 0$$

or $\quad \lambda^2 - 3\lambda - 10 = 0,$

i.e. $\quad (\lambda - 5)(\lambda + 2) = 0$

and hence, once again we obtain eigenvalues 5 and -2.

So, in this case **AB** and **BA** have the same eigenvalues.

Note: It would have been sufficient to show that **AB** and **BA** had the same characteristic equations.

4. **(i)** Using the result of Theorem 3, $\begin{bmatrix} 2 & 0 \\ 0 & 3 \end{bmatrix}$ has eigenvalues 2 and 3.

(ii) Using the result of Theorem 4, $\begin{bmatrix} 7 & 2 & 1 \\ 0 & 8 & 5 \\ 0 & 0 & 9 \end{bmatrix}$ has eigenvalues 7, 8 and 9.

5. The matrix **A** has eigenvalues given by

$$\begin{vmatrix} 2-\lambda & 3 \\ 3 & 2-\lambda \end{vmatrix} = 0$$

or $\quad (2-\lambda)^2 - 3^2 = 0.$

This gives eigenvalues 5 and -1. Substituting these values into $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, we obtain an eigenvector $[1 \quad 1]^T$ corresponding to the eigenvalue 5, and $[1 \quad -1]^T$ corresponding to the eigenvalue -1.

Thus Theorem 5 gives us that

$$\mathbf{P} = \begin{bmatrix} 1 & 1 \\ 1 & -1 \end{bmatrix} \quad \text{and } \mathbf{D} = \begin{bmatrix} 5 & 0 \\ 0 & -1 \end{bmatrix}$$

and you can check in this case that

$$\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}.$$

The matrix **P** is not unique, for it depends on the choice of eigenvectors. You should check that $\mathbf{P}^{-1}\mathbf{AP} = \mathbf{D}$ for your **P**. There are only two possible answers for **D**—the one above or

$$\mathbf{D} = \begin{bmatrix} -1 & 0 \\ 0 & 5 \end{bmatrix}$$

if the columns of **P** are reversed.

In other words, **D** must be a diagonal matrix with the eigenvalues of **A** on the diagonal.

6. We shall apply Procedure 3.3 to the matrix

$$\mathbf{A}_0 = \begin{bmatrix} 6 & 5 \\ 4 & 5 \end{bmatrix}$$

recording all numbers to three decimal places.

First iteration:

$$\mathbf{A}_0 = \mathbf{L}_0\mathbf{U}_0 = \begin{bmatrix} 1 & 0 \\ 0.667 & 1 \end{bmatrix}\begin{bmatrix} 6 & 5 \\ 0 & 1.667 \end{bmatrix}$$

$$\mathbf{A}_1 = \mathbf{U}_0\mathbf{L}_0 = \begin{bmatrix} 6 & 5 \\ 0 & 1.667 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0.667 & 1 \end{bmatrix} = \begin{bmatrix} 9.333 & 5 \\ 1.111 & 1.667 \end{bmatrix}$$

Second iteration:

$$\mathbf{A}_1 = \mathbf{L}_1\mathbf{U}_1 = \begin{bmatrix} 1 & 0 \\ 0.119 & 1 \end{bmatrix}\begin{bmatrix} 9.333 & 5 \\ 0 & 1.071 \end{bmatrix}$$

$$\mathbf{A}_2 = \mathbf{U}_1\mathbf{L}_1 = \begin{bmatrix} 9.333 & 5 \\ 0 & 1.071 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0.119 & 1 \end{bmatrix} = \begin{bmatrix} 9.929 & 5 \\ 0.128 & 1.071 \end{bmatrix}$$

Third iteration:

$$\mathbf{A}_2 = \mathbf{L}_2\mathbf{U}_2 = \begin{bmatrix} 1 & 0 \\ 0.013 & 1 \end{bmatrix}\begin{bmatrix} 9.929 & 5 \\ 0 & 1.007 \end{bmatrix}$$

$$\mathbf{A}_3 = \mathbf{U}_2\mathbf{L}_2 = \begin{bmatrix} 9.929 & 5 \\ 0 & 1.007 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0.013 & 1 \end{bmatrix} = \begin{bmatrix} 9.993 & 5 \\ 0.013 & 1.007 \end{bmatrix}.$$

Hence, an estimate for the eigenvalues of **A** would be 9.993 and 1.007 after 3 iterations.

You can check that the eigenvalues are in fact 10 and 1.

7. We start with the matrix

$$\mathbf{A}_0 = \begin{bmatrix} 7 & 5 \\ 3 & 5 \end{bmatrix}.$$

First iteration:

$$\mathbf{A}_0 = \mathbf{L}_0\mathbf{U}_0 = \begin{bmatrix} 1 & 0 \\ 0.429 & 1 \end{bmatrix}\begin{bmatrix} 7 & 5 \\ 0 & 2.857 \end{bmatrix}$$

$$\mathbf{A}_1 = \mathbf{U}_0\mathbf{L}_0 = \begin{bmatrix} 7 & 5 \\ 0 & 2.857 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0.429 & 1 \end{bmatrix} = \begin{bmatrix} 9.143 & 5 \\ 1.224 & 2.857 \end{bmatrix}$$

Second iteration:

$$\mathbf{A}_1 = \mathbf{L}_1\mathbf{U}_1 = \begin{bmatrix} 1 & 0 \\ 0.134 & 1 \end{bmatrix}\begin{bmatrix} 9.143 & 5 \\ 0 & 2.188 \end{bmatrix}$$

$$\mathbf{A}_2 = \mathbf{U}_1\mathbf{L}_1 = \begin{bmatrix} 9.143 & 5 \\ 0 & 2.188 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0.134 & 1 \end{bmatrix} = \begin{bmatrix} 9.813 & 5 \\ 0.293 & 2.188 \end{bmatrix}$$

Third iteration:

$$\mathbf{A}_2 = \mathbf{L}_2\mathbf{U}_2 = \begin{bmatrix} 1 & 0 \\ 0.0299 & 1 \end{bmatrix}\begin{bmatrix} 9.813 & 5 \\ 0 & 2.038 \end{bmatrix}$$

$$\mathbf{A}_3 = \mathbf{U}_2\mathbf{L}_2 = \begin{bmatrix} 9.813 & 5 \\ 0 & 2.038 \end{bmatrix}\begin{bmatrix} 1 & 0 \\ 0.0299 & 1 \end{bmatrix} = \begin{bmatrix} 9.962 & 5 \\ 0.0609 & 2.038 \end{bmatrix}.$$

Hence, an approximation for the eigenvalues of $\mathbf{A}_0$ after three iterations would be 9.962 and 2.038.

Solutions to the exercises in Section 4

The answers to most of the exercises can be checked using the 'black box' routine (Option 23). However, Exercises 1 and 3 require the additional explanations given below.

1. (i) Starting with the initial vector $\mathbf{y}_0 = [3 \quad 1]^T$ (Case (a)) the scheme produces the eigenvalue (and corresponding eigenvector) of largest modulus, as expected. However, starting with the initial vector $\mathbf{y}_0 = [-2 \quad 1]^T$ (Case (b)) the scheme produces the eigenvalue of *smallest* modulus. This unexpected result can be explained by observing that the scheme stipulates that the initial vector $\mathbf{y}_0$ must not be an eigenvector. In Case (b) the initial vector was chosen to be $\mathbf{y}_0 = [-2 \quad 1]^T$, which is an eigenvector corresponding to the eigenvalue of smallest modulus.

You might well ask how we can avoid choosing $\mathbf{y}_0$ to be an eigenvector. If the scheme reaches the solution in only one or two iterations, you have reason to be very suspicious!

(ii) The initial vectors $\mathbf{y}_0$—however close they are to an eigenvector corresponding to the eigenvalue of smallest modulus—are not in fact eigenvectors and do, as the geometry on the television programme suggests, get dragged round to the dominant eigenvector.

3. (i) The answers differ slightly from those in Exercise 2 because the LR method stops when the magnitude of all the elements below the diagonal in $\mathbf{A}_r$ are less than the tolerance you stated. This does not in fact tell you how accurate your answer is going to be. It just ensures reasonable accuracy in the answer you give.

Solutions to the end of unit test

1. $\det(\mathbf{A} - \lambda\mathbf{I}) = 0$ gives

$$\begin{vmatrix} 7-\lambda & 14 \\ 3 & 8-\lambda \end{vmatrix} = 0.$$

Hence,

$$\begin{aligned} &(7-\lambda)(8-\lambda) - 42 = 0 \\ \text{or} \quad &\lambda^2 - 15\lambda + 14 = 0 \\ \text{i.e.} \quad &(\lambda - 14)(\lambda - 1) = 0. \end{aligned}$$

So, the eigenvalues of $\mathbf{A}$ are 1 and 14. The corresponding eigenvectors can be found by solving $(\mathbf{A} - \lambda\mathbf{I})\mathbf{x} = \mathbf{0}$, putting λ equal to each eigenvalue in turn.

Case $\lambda = 1$: we get the equations

$$\begin{aligned} 6x_1 + 14x_2 &= 0 \\ 3x_1 + 7x_2 &= 0. \end{aligned}$$

Thus we obtain a solution $x_2 = k$, $x_1 = -\frac{7}{3}k$, and hence an eigenvector $[-7 \quad 3]^T$ corresponding to the eigenvalue 1.

Case $\lambda = 14$: we get the equations

$$\begin{aligned} -7x_1 + 14x_2 &= 0 \\ 3x_1 - 6x_2 &= 0. \end{aligned}$$

Thus we obtain a solution $x_2 = k$, $x_1 = 2k$, and hence an eigenvector $[2 \quad 1]^T$ corresponding to the eigenvalue 14.

2. (i) To find the eigenvalue of largest modulus, we use direct iteration with scaling.

Step 1:

$$\mathbf{y}_1' = \begin{bmatrix} 7 & 14 \\ 3 & 8 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 21 \\ 11 \end{bmatrix}, \qquad \alpha_1 = 21.$$

So, dividing by 21 we obtain $\mathbf{y}_1 = [1 \quad 0.524]^T$.

Step 2:

$$\mathbf{y}_2' = \begin{bmatrix} 7 & 14 \\ 3 & 8 \end{bmatrix}\begin{bmatrix} 1 \\ 0.524 \end{bmatrix} = \begin{bmatrix} 14.333 \\ 7.190 \end{bmatrix}, \qquad \alpha_2 = 14.333.$$

So, dividing by 14.333 we obtain $\mathbf{y}_2 \doteq [1 \quad 0.502]^T$.

Step 3:

$$\mathbf{y}_3' = \begin{bmatrix} 7 & 14 \\ 3 & 8 \end{bmatrix}\begin{bmatrix} 1 \\ 0.502 \end{bmatrix} = \begin{bmatrix} 14.023 \\ 7.013 \end{bmatrix}, \qquad \alpha_3 = 14.023.$$

So, dividing by 14.023 we obtain $\mathbf{y}_3 = [1 \quad 0.500]^T$.

As $\mathbf{y}_2$ and $\mathbf{y}_3$ agree to 2 significant figures, we shall stop the iteration. Hence, to two significant figures the eigenvalue of largest modulus is 14, and a corresponding eigenvector is $[1 \quad 0.50]^T$.

(ii) To find the eigenvalue of smallest modulus, we use inverse iteration $\mathbf{y}_{r+1}' = \mathbf{A}^{-1}\mathbf{y}_r$, where

$$\mathbf{A}^{-1} = \frac{1}{14}\begin{bmatrix} 8 & -14 \\ -3 & 7 \end{bmatrix}.$$

Step 1:

$$\mathbf{y}_1' = \frac{1}{14}\begin{bmatrix} 8 & -14 \\ -3 & 7 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} -0.429 \\ 0.286 \end{bmatrix}, \qquad \alpha_1 = -0.429.$$

So, dividing by -0.429 we obtain $\mathbf{y}_1 = [1 \quad -0.667]^T$.

Step 2:

$$\mathbf{y}_2' = \frac{1}{14}\begin{bmatrix} 8 & -14 \\ -3 & 7 \end{bmatrix}\begin{bmatrix} 1 \\ -0.667 \end{bmatrix} = \begin{bmatrix} 1.238 \\ -0.548 \end{bmatrix}, \qquad \alpha_2 = 1.238.$$

So, dividing by 1.238 we obtain $\mathbf{y}_2 = [1 \quad -0.442]^T$.

Step 3:

$$\mathbf{y}_3' = \frac{1}{14}\begin{bmatrix} 8 & -14 \\ -3 & 7 \end{bmatrix}\begin{bmatrix} 1 \\ -0.442 \end{bmatrix} = \begin{bmatrix} 1.014 \\ -0.435 \end{bmatrix}, \qquad \alpha_3 = 1.014.$$

So, dividing by 1.014 we obtain $\mathbf{y}_3 = [1 \quad -0.430]^T$.

Step 4:

$$\mathbf{y}_4' = \frac{1}{14}\begin{bmatrix} 8 & -14 \\ -3 & 7 \end{bmatrix}\begin{bmatrix} 1 \\ -0.430 \end{bmatrix} = \begin{bmatrix} 1.001 \\ -0.430 \end{bmatrix}, \qquad \alpha_4 = 1.001.$$

So, dividing by 1.001 we obtain $\mathbf{y}_4 = [1 \quad -0.430]^T$.

Now, $\mathbf{y}_4$ and $\mathbf{y}_3$ agree to 2 significant figures and so we shall stop the iteration. Hence, an approximation for the eigenvalue of smallest modulus of $\mathbf{A}$ is $1/1.001 = 1.0$ (to 2 significant figures). The corresponding eigenvector is $[1 \quad -0.430]^T$.

(iii) $\mathbf{A} - 12\mathbf{I} = \begin{bmatrix} -5 & 14 \\ 3 & -4 \end{bmatrix}$, so

$(\mathbf{A} - 12\mathbf{I})^{-1} = \frac{1}{22}\begin{bmatrix} 4 & 14 \\ 3 & 5 \end{bmatrix}$. Inverse iteration

$\mathbf{y}_{r+1}' = (\mathbf{A} - 12\mathbf{I})^{-1}\mathbf{y}_r$ gives

Step 1:

$$\mathbf{y}_1' = \frac{1}{22}\begin{bmatrix} 4 & 14 \\ 3 & 5 \end{bmatrix}\begin{bmatrix} 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 0.818 \\ 0.364 \end{bmatrix}, \qquad \alpha_1 = 0.818.$$

So, dividing by 0.818 we obtain $\mathbf{y}_1 = [1 \quad 0.444]^T$.

Step 2:

$$\mathbf{y}_2' = \frac{1}{22}\begin{bmatrix}4 & 14\\3 & 5\end{bmatrix}\begin{bmatrix}1\\0.444\end{bmatrix}$$

$$= \begin{bmatrix}0.465\\0.237\end{bmatrix}, \quad \alpha_2 = 0.465.$$

So, dividing by 0.465 we obtain $\mathbf{y}_2 = [1 \quad 0.511]^T$.

Step 3:

$$\mathbf{y}_3' = \frac{1}{22}\begin{bmatrix}4 & 14\\3 & 5\end{bmatrix}\begin{bmatrix}1\\0.511\end{bmatrix}$$

$$= \begin{bmatrix}0.507\\0.252\end{bmatrix}, \quad \alpha_3 = 0.507.$$

So, dividing by 0.507 we obtain $\mathbf{y}_3 = [1 \quad 0.498]^T$.

Step 4:

$$\mathbf{y}_4' = \frac{1}{22}\begin{bmatrix}4 & 14\\3 & 5\end{bmatrix}\begin{bmatrix}1\\0.498\end{bmatrix}$$

$$= \begin{bmatrix}0.499\\0.250\end{bmatrix}, \quad \alpha_4 = 0.499.$$

So, dividing by 0.499 we obtain $\mathbf{y}_4 = [1 \quad 0.500]^T$.

Now, $\mathbf{y}_3$ and $\mathbf{y}_4$ agree to 2 significant figures, and so we shall stop the iteration. Hence, an approximation for the eigenvalue nearest to 12 is $\dfrac{1}{0.499} + 12 = 14$ (to 2 significant figures). The corresponding eigenvalue is $[1 \quad 0.5]^T$.

3. First iteration:

$$\mathbf{A}_0 = \mathbf{L}_0\mathbf{U}_0 = \begin{bmatrix}1 & 0\\0.429 & 1\end{bmatrix}\begin{bmatrix}7 & 14\\0 & 2\end{bmatrix}$$

$$\mathbf{A}_1 = \mathbf{U}_0\mathbf{L}_0 = \begin{bmatrix}7 & 14\\0 & 2\end{bmatrix}\begin{bmatrix}1 & 0\\0.429 & 1\end{bmatrix}$$

$$= \begin{bmatrix}13.000 & 14\\0.857 & 2\end{bmatrix}$$

Second iteration:

$$\mathbf{A}_1 = \mathbf{L}_1\mathbf{U}_1 = \begin{bmatrix}1 & 0\\0.0659 & 1\end{bmatrix}\begin{bmatrix}13 & 14\\0 & 1.077\end{bmatrix}$$

$$\mathbf{A}_2 = \mathbf{U}_1\mathbf{L}_1 = \begin{bmatrix}13 & 14\\0 & 1.077\end{bmatrix}\begin{bmatrix}1 & 0\\0.0659 & 1\end{bmatrix}$$

$$= \begin{bmatrix}13.923 & 14\\0.07100 & 1.077\end{bmatrix}$$

Third iteration:

$$\mathbf{A}_2 = \mathbf{L}_2\mathbf{U}_2 = \begin{bmatrix}1 & 0\\0.00510 & 1\end{bmatrix}\begin{bmatrix}13.923 & 14\\0 & 1.006\end{bmatrix}$$

$$\mathbf{A}_3 = \mathbf{U}_2\mathbf{L}_2 = \begin{bmatrix}13.923 & 14\\0 & 1.006\end{bmatrix}\begin{bmatrix}1 & 0\\0.00510 & 1\end{bmatrix}$$

$$= \begin{bmatrix}13.994 & 14\\0.00513 & 1.006\end{bmatrix}.$$

Thus, approximations for the eigenvalues of $\mathbf{A}_0$ are 13.994 and 1.006—or 14 and 1.0 to two significant figures.

4. We know that

$$\mathbf{A}\mathbf{x}_i = \lambda_i\mathbf{x}_i \qquad \text{for } i = 1, 2, \ldots, n \tag{1}$$

(i) Multiplying Equation (1) by $\mathbf{A}$, we get

$$\mathbf{A}^2\mathbf{x}_i = \lambda_i\mathbf{A}\mathbf{x}_i$$
$$= \lambda_i(\lambda_i\mathbf{x}_i) \qquad \text{(using (1))}.$$

So $\mathbf{A}^2\mathbf{x}_i = \lambda_i^2\mathbf{x}_i$ (2)

Multiplying Equation (2) by $\mathbf{A}$, we get

$$\mathbf{A}^3\mathbf{x}_i = \lambda_i^2\mathbf{A}\mathbf{x}_i$$
$$= \lambda_i^2(\lambda_i\mathbf{x}_i) \qquad \text{(from (1))}.$$

So $\mathbf{A}^3\mathbf{x}_i = \lambda_i^3\mathbf{x}_i$.

Thus, the eigenvalues of $\mathbf{A}^3$ are $\lambda_1^3, \lambda_2^3, \ldots, \lambda_n^3$, and the corresponding eigenvectors are the same as those of $\mathbf{A}$, i.e. $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$.

(ii) Adding $p\mathbf{I}\mathbf{x}_i$ (or equivalently $p\mathbf{x}_i$) to both sides of (1), we get

$$(\mathbf{A} + p\mathbf{I})\mathbf{x}_i = (\lambda_i + p)\mathbf{x}_i. \tag{3}$$

Similarly, subtracting $p\mathbf{I}\mathbf{x}_i$ (or equivalently $p\mathbf{x}_i$) from both sides of (1), we get

$$(\mathbf{A} - p\mathbf{I})\mathbf{x}_i = (\lambda_i - p)\mathbf{x}_i. \tag{4}$$

Re-arranging (3) and (4) we get

$$\mathbf{x}_i = \frac{(\mathbf{A} + p\mathbf{I})\mathbf{x}_i}{\lambda_i + p} \qquad \text{and} \qquad \mathbf{x}_i = \frac{(\mathbf{A} - p\mathbf{I})\mathbf{x}_i}{\lambda_i - p}.$$

Equating these two, we get

$$\frac{(\mathbf{A} - p\mathbf{I})\mathbf{x}_i}{\lambda_i - p} = \frac{(\mathbf{A} + p\mathbf{I})\mathbf{x}_i}{\lambda_i + p}.$$

Left-multiplying this relation by $(\mathbf{A} + p\mathbf{I})^{-1}(\lambda_i - p)$ we get

$$(\mathbf{A} + p\mathbf{I})^{-1}(\mathbf{A} - p\mathbf{I})\mathbf{x}_i = \frac{\lambda_i - p}{\lambda_i + p}\mathbf{x}_i.$$

So, the eigenvalues of $(\mathbf{A} + p\mathbf{I})^{-1}(\mathbf{A} - p\mathbf{I})$ are $\dfrac{\lambda_i - p}{\lambda_i + p}$ for $i = 1, 2, \ldots, n$.

The eigenvectors are the same as those of $\mathbf{A}$, i.e. $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_n$. (You can probably see by now the connection between the eigenvalues of $\mathbf{A}$ and those of various 'functions' of $\mathbf{A}$.)

5. (i)

$$\mathbf{L}^T\mathbf{L} = \begin{bmatrix}\sqrt{a} & \dfrac{b}{\sqrt{a}}\\ 0 & \sqrt{\dfrac{\det\mathbf{A}}{a}}\end{bmatrix}\begin{bmatrix}\sqrt{a} & 0\\ \dfrac{b}{\sqrt{a}} & \sqrt{\dfrac{\det\mathbf{A}}{a}}\end{bmatrix}$$

$$= \begin{bmatrix}a + \dfrac{b^2}{a} & \dfrac{b}{a}\sqrt{\det\mathbf{A}}\\ \dfrac{b}{a}\sqrt{\det\mathbf{A}} & \dfrac{1}{a}\det\mathbf{A}\end{bmatrix}$$

which is a symmetric matrix.

(ii) From the given decomposition:

$$\begin{bmatrix}9 & 8\\8 & 9\end{bmatrix} = \begin{bmatrix}3 & 0\\2.667 & 1.374\end{bmatrix}\begin{bmatrix}3 & 2.667\\0 & 1.374\end{bmatrix}.$$

6. (i) We have

$$\mathbf{A}_0 = \mathbf{L}_0\mathbf{L}_0^T$$
$$\mathbf{A}_1 = \mathbf{L}_0^T\mathbf{L}_0.$$

We know from Theorem 1 of Section 3 that, for any square matrices $\mathbf{A}$ and $\mathbf{B}$, $\mathbf{AB}$ and $\mathbf{BA}$ have the same eigenvalues and so $\mathbf{A}_0$ and $\mathbf{A}_1$ have the same eigenvalues. Similarly, so do $\mathbf{A}_1$ and $\mathbf{A}_2$ etc. Thus the sequence $\mathbf{A}_0, \mathbf{A}_1, \mathbf{A}_2, \ldots$ all have the same eigenvalues.

(ii) To cut down the working, we shall use the result of Exercise 5(i) which tells us that if $\mathbf{A}_r = \begin{bmatrix}a & b\\b & c\end{bmatrix}$ then

$$\mathbf{A}_{r+1} = \begin{bmatrix}a + \dfrac{b^2}{a} & \dfrac{b}{a}\sqrt{\det\mathbf{A}}\\ \dfrac{b}{a}\sqrt{\det\mathbf{A}} & \dfrac{1}{a}\det\mathbf{A}\end{bmatrix}.$$

First iteration: for $\mathbf{A}_0$, $a = 9$, $b = 8$, $\det \mathbf{A} = 17$, and so

$$\mathbf{A}_1 = \begin{bmatrix} 9 + \frac{64}{9} & \frac{8}{9}\sqrt{17} \\ \frac{8}{9}\sqrt{17} & \frac{1}{9} \times 17 \end{bmatrix} = \begin{bmatrix} 16.111 & 3.6650 \\ 3.6650 & 1.8889 \end{bmatrix}.$$

Second iteration: $a = 16.111$, $b = 3.6650$, $\det \mathbf{A} = 16.9998$, and so

$$\mathbf{A}_2 = \begin{bmatrix} 16.111 + \dfrac{(3.6650)^2}{16.111} & \dfrac{3.6650 \times \sqrt{16.9998}}{16.111} \\ \dfrac{3.6650 \times \sqrt{16.9998}}{16.111} & \dfrac{16.9998}{16.111} \end{bmatrix} = \begin{bmatrix} 16.9447 & 0.93794 \\ 0.93794 & 1.05517 \end{bmatrix}.$$

Third iteration: $a = 16.9447$, $b = 0.93794$, $\det \mathbf{A} = 16.9998$, and so

$$\mathbf{A}_3 = \begin{bmatrix} 16.9447 + \dfrac{(0.93794)^2}{16.9447} & 0.93794 \times \dfrac{\sqrt{16.9998}}{16.9447} \\ 0.93794 \times \dfrac{\sqrt{16.9998}}{16.9447} & \dfrac{16.9998}{16.9447} \end{bmatrix} = \begin{bmatrix} 16.9966 & 0.2282 \\ 0.2282 & 1.0033 \end{bmatrix}.$$

As we know that $\mathbf{A}_0$, $\mathbf{A}_1$, $\mathbf{A}_2$ and $\mathbf{A}_3$ all have the same eigenvalues, we can get an approximation from $\mathbf{A}_3$ for the eigenvalues of $\mathbf{A}_0$ of 16.9966 and 1.0033—or 17 and 1.0 to two significant figures.

7. Given that both methods work, it would be advisable to choose the $\mathbf{LL}^T$ decomposition, simply because there are less elements to find—and hence, less work involved.

The Open University
0 335 14050 5